NATIONAL QUALITY FORUM

+ + + + +

PATIENT SAFETY 2015-2017
STEERING COMMITTEE

+ + + + +

WEDNESDAY
JULY 27, 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.
LYNDA SMIRZ, MD, MBA, Universal Health Systems of Delaware
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:
ANDREW ANDERSON, MHA, Senior Project Manager
KAREN JOHNSON, MS, Senior Director
ANDREW LYZENGA, MPP, Senior Director
ELISA MUNTHALI, MPH, Vice President, Quality Measurement
JESSE PINES, MD, Senior Director
DESMIRRA QUINNONEZ, Project Analyst
MARCIA WILSON, PhD, MBA, Senior Vice President, Quality Measurement

ALSO PRESENT:

KRISTEN BUTTERFIELD, MPH, Pharmacy Quality Alliance (PQA)
DEL CONYERS, National PACE Association
NANCY DUNTON, PhD, FAAN, University of Kansas School of Nursing
WOODY EISENBERG, MD, Pharmacy Quality Alliance (PQA)
ERIN GIAVONETTI, PhD, National Committee for Quality Assurance (NCQA)
TAMIKA GLADNEY, Centers for Medicare and Medicaid Services (CMS)
LISA HINES, PharmD, Pharmacy Quality Alliance (PQA)*
ROBYN McGONIGAL, MD, MPH, Kidney Care Quality Alliance (KCQA)
EMILY MORDEN, MSW, National Committee for Quality Assurance (NCQA)*
ROBYN NISHIMI, PhD, Kidney Care Quality Alliance (KCQA)
LYNN PEZZULLO, RPh, CPEHR, Pharmacy Quality Alliance (PQA)
BOB REHM, MBA, National Committee for Quality Assurance (NCQA)
MARK STEWART, MPH, Econometrica, Inc.

* present by teleconference

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

Consideration of Candidate Meaures (cont.)

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

(9:00 a.m.)

CO-CHAIR SEPTIMUS: Okay. We're going to go ahead and get started. So we want to start on time. We want to be focused and efficient. Those of you who have looked at the agenda and have seen the number of measures that we're considering will know that this is considerably shorter than last year. And there's no major controversial issues like sepsis and PSI 90.

CO-CHAIR THRAEN: So we need to have some fun as well.

CO-CHAIR SEPTIMUS: We need to have some fun, right. But nonetheless, we really have the opportunity that when we finish here tomorrow at 3 o'clock to be finished and not have to have a follow up call if we stay focused.

You'll also notice on the agenda that there are probably more new measures than there are maintenance measures. So that's another sort of challenge that we have. And we have some great new

measures to discuss, so I hope we'll stay focused.

The other thing I noticed last night as I passed through the lobby several times is seeing some of you there. And one of the wonderful things about being a member of this committee is getting to know each and every one of you and knowing the incredible wealth of knowledge and wisdom that you have brought to this process. And I'm sure Iona will also say this, but not only --

CO-CHAIR THRAEN: So don't say it.

CO-CHAIR SEPTIMUS: No, wait a minute. Not knowing we come together to do the work that we think is very important, but I think we've actually bonded as a team. And I think that's an incredibly wonderful experience beyond the satisfaction of the work that we do.

So I think we all can say we learn a tremendous amount from each other and we come away actually being better for it. So we thank you because we know this is an incredible amount of work.

And last year, as you know, we did a

record number of measures. CSAC actually approved our agenda with no comment and it passed through. Now, I don't know if that's ever happened before.

PARTICIPANT: It may not have.

CO-CHAIR SEPTIMUS: I don't think that's ever happened before, so the credit for that goes to all of you for the incredible amount of hard work we put in last year.

And it was, make no mistakes about it, it was a lot of work. The year before was a lot of work, but I think last year was especially challenging because of the number of measures that we had.

So this year we have a very manageable number, but a lot of new measures to consider and a number of eMeasures as well. We sort of got introduced to eMeasures last year for the first time.

So we're looking forward to a great day. Stay focused. Let's stay on time. And we have dinner tonight so we can -- and as has been our tradition, the chair will buy wine for everybody,

so it's something to think about. But the condition is we have to finish the agenda for today. Okay? With that I'm going to turn it over to Iona.

CO-CHAIR THRAEN: I don't have anything else to say other than welcome and, hopefully, we'll have some fun along the way. I know that you guys put a lot of work into this and so thank you. That's it.

DR. WILSON: Good morning. My name's Marcia Wilson. I'm senior vice president here at the Quality Measurement Department and I am going to fill in for our legal counsel, Ann Hammersmith, and do the disclosures of interest.

So you all know when you were appointed to this committee you filled out a form where we asked you a lot of questions. And today we do an oral disclosure of interest and we combine it with the introductions.

So when you do your disclosure it is not necessary to summarize your resume. We already know what an incredible group of people you are, and as Ed said you bring a wealth of experience to

the table.

But we are interested in your disclosing any work that is directly related to the issues and the measures before the committee today.

This could be grants for research and it's not limited to activities where you get paid because you may serve on a board. That is something also that you would like to disclose. And again, the activities that you need to disclose are those related to this subject matter.

Just a couple of reminders, you do sit on this committee as an individual even though you all come from organizations, and you don't represent the interests of your employer. And just because you disclose, it does not mean you have a conflict of interest. But we do these verbal disclosures in the spirit of transparency because, of course, we're all about transparency here at NQF.

So we'll go around the room, first of all, and then I'll turn to some of the committee members who are on the phone today, and if you would

state your name, who you're with and if you have any activities that you need to disclose.

And I think we've already -- are aware of a couple of conflicts where we have recusals, so we've made note of those, but you can state those again.

And I'll go ahead and start with our co-chairs, Iona?

CO-CHAIR THRAEN: Iona Thraen. I'm the director of patient safety for the Utah Department of Health. I have not participated in any measure development.

CO-CHAIR SEPTIMUS: Ed Septimus, medical director of infection prevention and epidemiology at HCA in Houston and also a professor of internal medicine at Texas A&M College of Medicine. And I'm obligated to say howdy.

DR. WILSON: And we can turn to Lisa McGiffert, if you want to go ahead.

MEMBER MCGIFFERT: That was fast. I'm Lisa McGiffert with Consumer's Unions Safety Patient Project and I don't have any conflicts to

disclose.

MEMBER APPLEGATE: Kimberly Applegate and I'm a professor of pediatric radiology at Emory University and I have no conflict of interest.

MEMBER SCHULTZ: Leslie Schultz. I'm with Premier, Inc. I'm with the Safety Institute and I have nothing to disclose.

MEMBER DANFORTH: Missy Danforth, I'm the vice president for hospital ratings at The Leapfrog Group and I have nothing to disclose.

MEMBER COOK: Hi, I'm Chris Cook. With bioMerieux and I have no conflicts to disclose.

MEMBER YU: I'm Yanling Yu. I'm a research scientist and then also a patient advocate. I have no conflict of interest to disclose.

CO-CHAIR SEPTIMUS: I'm sorry. I forgot to mention I have no conflicts. I'm sorry.

MEMBER WU: We were wondering. Albert Wu, I'm an internist and professor at the Johns Hopkins Bloomberg School of Public Health and part

of the Armstrong Institute for Patient Safety, no conflicts.

MEMBER DEED: I'm -- oh, that's helpful. Yes. Lesson learned, maybe. I'm Martha Deed. I'm a patient advocate and I'm here on behalf of having been nominated by name. And I have no conflicts of interest.

MEMBER WANG: Hi. Good morning. I'm Tracy Wang, program director of community health initiatives with Anthem and I have no disclosures.

MEMBER EDELSTEIN: Good morning, Theresa Edelstein. I'm vice president of post-acute care policy at the New Jersey Hospital Association. My conflict to disclose is I'm a member of the PACE technical expert panel.

MEMBER ARDIZZONE: I'm Laura Ardizzone from the American Association of Nurse Anesthetists. I'm also the director of nurse anesthesia services as Memorial Sloan Kettering Cancer Center and I have no disclosures.

MEMBER GELINAS: Good morning, everyone. I'm Lillee Gelinas with CHRISTUS

Health. I'm also editor-in-chief of American Nurse Today.

I have two major disclosures. First is that I co-chair with Dr. Mary Naylor, the NQF Nursing-Sensitive Measures Committee which developed several of the nursing-sensitive measures now in use in the U.S.

And secondly, I'm a member of the ANA Tipping Point Committee charged with eMeasure development beginning with pressure ulcers. And therefore I'm a primary investigator for the eMeasure pressure ulcer work across all of CHRISTUS Health.

MEMBER QUIGLEY: Good morning. I'm Pat Quigley and I come to you as Patricia A. Quiqley, nurse consultant. So that's who I am with because I retired from the Department of Veterans Affairs February 1. So I'll have to update my bio. And colleagues, I have nothing to disclose.

MEMBER ALEXANDER: I'm Charlotte Alexander with Memorial Hermann and I have nothing to disclose.

MEMBER WEBB: I am Kendall Webb. I work at the UF Health Jacksonville facility as the CMIO. And I'm here actually with American College of Emergency Physicians out of their Quality and Performance Committee. I have nothing to disclose.

MEMBER LAWLESS: I'm Dr. Steve Lawless. I'm the senior vice president and chief clinical officer for the Nemours System and I have nothing to disclose.

DR. WILSON: Thank you. And now we'll go to any of the committee members who are on the phone. I think Michelle Schreiber is on the phone.

MEMBER SCHREIBER: Yes. Thank you. Good morning, I'm Michelle Schreiber. I'm the chief quality officer of the Henry Ford Health System in Detroit and I have nothing to disclose.

DR. WILSON: Thank you. Are any other committee members on the phone with us at this time? Okay. Thank you. We may have a committee member joining us a little later on and when they do, they can do the disclosure when they come in.

Thank you all for that information. And I'd like to remind you that if you believe at any time during this discussion today or tomorrow either you think of something that might be a conflict or someone else says something that you think might be a conflict, please don't remain silent. You can approach either your co-chairs or any of the NQF staff if you have concerns.

We would much rather have you bring something forward so we could discuss it rather than sit and be concerned that there was any type of conflict going on.

Based on what you've heard from your colleagues around the committee, do you have any questions at this time for me?

CO-CHAIR THRAEN: I have one.

DR. WILSON: Yes, Iona.

CO-CHAIR THRAEN: Okay. So Victoria Rich was intending to be a presenter on the PACE-acquired pressure ulcer injury prevent prevalence rate.

DR. WILSON: Uh-huh.

CO-CHAIR THRAEN: She's had a family emergency.

DR. WILSON: Right.

CO-CHAIR THRAEN: She's not going to be here. We were going to ask Susan Moffatt. She's not joined us quite yet. And then Chris is the back up. And Lillee had volunteered to present, but because she was on the PACE group is that a conflict of interest?

DR. WILSON: Yes, it would.

CO-CHAIR THRAEN: Okay.

DR. WILSON: We would not have her present.

CO-CHAIR THRAEN: So Chris, you're going to be on board to do the presentation on that particular measure?

DR. WILSON: So Chris is on deck for that one.

CO-CHAIR THRAEN: Okay. Thank you.

DR. WILSON: Yes. Any other questions? Yes.

MEMBER GELINAS: First of all, thank

you to the NQF staff for individual microphones. I guess many of us who have been a part of NQF for some time remember sharing them, so appreciate that.

As a part of protocol would you remind us if we do have to recuse ourselves limitations on discussion versus limitations on actually voting?

DR. WILSON: Thank you for that question. If you are recused you may not join in the discussion. You do not have to leave the room, but may not participate in the discussion nor can you vote. So that's our policy any time someone has a recusal. You can listen, but you may not speak or vote.

CO-CHAIR SEPTIMUS: Thanks. Thanks for that clarification. Appreciate that. Any other questions of Marsha? Okay. Well, that's great. So we have Andrew and Andrew here, so.

CO-CHAIR THRAEN: Andrew squared.

CO-CHAIR SEPTIMUS: Yes. Andrew Anderson, have you been to one of our meetings

before?

MR. ANDERSON: Yes, I was.

CO-CHAIR THRAEN: Yes, he was here last one.

CO-CHAIR SEPTIMUS: Okay. I'm sorry. I didn't -- okay. So, what --

CO-CHAIR THRAEN: One's Drew and one's Andrew. He's Drew.

CO-CHAIR SEPTIMUS: I told you we're going to have fun, right? So we're ahead of schedule, so that's good. So I'm going to turn it over to Drew and Andrew. Get that right now? See, I am a learner.

Talk about product introduction and overview of the evaluation process and then, I think we'll just go right into talking about the patient's safety measure portfolio that we're going to review. And I know that we're going to mention about certain new processes in terms of how we can discuss maintenance measures. So I'll turn it over to both of you.

MR. ANDERSON: Sure. So I'll get

started just with some housekeeping items. Let's see. So just as a reminder, we're going to have two breaks today, one for lunch and then one at 3:30. Depending on how we move through the measures we might adjust that a little bit.

And then we've been streaming the web information if you want to log in for wi-fi. We have a couple documents on SharePoint that you might want to pull up if you have your laptop with you.

We posted last night a related and competing comparison table. It's now on the main part of the committee SharePoint, so if you could pull that up once we get to that section, you can follow along. But we'll also be pulling it up here as we discuss the measures. And as you know the bathrooms are out and you can leave at any time, outside.

I'm just going to skip -- how does this work? Okay. So as you all are familiar these are just some ground rules for the meeting. You all have reviewed the measures beforehand already. We

want to make sure that we're basing all of your evaluations on the measure criteria.

We've including all of the measure evaluation forms with the folders that we passed out. So if you didn't get a chance to look at those, those should be at your place setting and we can get you that if you haven't been given it.

The other thing is if you can make sure that you stay in the room at all times unless you're at the restroom, keeping your comments concise and focused, try not to repeat too many things if you can avoid it, and then also allowing everyone to contribute. As you remember from last year, if you want to speak just put up your card and one of the co-chairs will call on you. Okay.

You all are already familiar with this, but we have a lot of public participants in the room and some on the line. Where we are in the eight-step consensus development process is the standards review.

And here, the measure evaluation criteria that you all are already familiar with,

for each measure we'll be walking through the importance to measure and report, talking about performance gaps.

We'll look at the scientific acceptability of the measures and measure properties, so the reliability and validity of the measures, assessing whether or not the measures are feasible, usable. And then we'll also be looking at that related and competing table to make sure that we choose the best in class if there are any measures that are competing with each other.

So just a quick overview of what we're looking at today. I'm just going to quickly go through some of our previous work that we've done in this area for patient safety. This is the third cycle of this project. So you all are very familiar, but, again, for public attendees. And then I'll turn it over to Andrew in a moment to go over the portfolio.

We've almost gotten to the point where we've reviewed almost all of our maintenance measures in this project, so later in this meeting

we'll be talking about gaps in measurements. And we hope to spend a little bit of time talking about where we can advance measurement in this area.

I know last year, of course, we were always talking about developing more outcome measures, but getting more detail around that and then having a Q and A discussion.

So like I said, this is the third cycle of patient safety. This is one of our longest consensus development projects. We've been doing patient safety projects almost since the beginning since NQF started its work.

We have a number of other projects that this patient safety overlaps with because it's a cross cutting area. And some of them are safe and better practices for better healthcare that came out in 2010, our report on serious reportable events and our common formats project that's been going on for some time.

And just as a reminder NQF has our measure applications partnership and that's why NQF decides while we have a number of workgroups,

clinician workgroups and a few others, that come up with recommendations for measures for federal programs. And we've been having a lot of conversations about intended use.

Patient safety is among one of MAP's family of measure projects and it was one of the frameworks that was originally developed, like I mentioned earlier.

And then we also have the National Quality Partners which convenes action teams, some of them around maternity care, re-admissions, patients and family engagement. NQP is also looking at an action team around shared decision-making. So a number of other things that go in and tie into this work. Okay.

So I'm going to turn it over to Andrew to go over the measure portfolio.

CO-CHAIR SEPTIMUS: Just one quick -- I may have missed it, but did you mention the measurement incubator work? If any of you want to make mention, not because I sit on that group, but I just think it's an important activity that the

committee may want to hear about.

MR. LYZENGA: Yes, that's a good point, Ed. And the last couple years we've been working to sort of develop a program that we're calling the Measure Incubator.

NQF hasn't previously been involved at all in the development of measures and we're really still not involved in development, but we're sort of working to create an environment wherein, you know, measured development can be advanced and incubated, so to speak.

What we're doing is trying to serve as a matchmaker of sorts, bringing together the folks who have a good measure concept or idea, people or groups who have expertise in measure development, maybe groups that have funding and have an interest in advancing measurement in a particular area, and just any sort of group or person or people who can contribute to development of new and innovative measures in a particular area.

We're going to try to bring them together and help, again, incubate and sort of

accelerate the development of those measures in some key gap areas where we're really in need of measures.

So any questions about that would be welcomed as well, but that's a pretty exciting new area we're getting into. I don't know if you want to add anything to that, Ed.

CO-CHAIR SEPTIMUS: No, that we're just really exploring where this should go. Anyone who's ever been involved in measure development, you can see the amount of work it takes to bring a measure forward to go through the rigorous process.

There's also a fair amount of expense involved in doing it. So this is a way to help facilitate new measures that we think, or not we, but the community think should be developed and how can they get them developed in a way that meets the rigorous standards that it needs to make to get through the NQF process.

So I think it's a really important step forward. It's a baby step, but it's an important

step forward. Before I forget, we need to congratulate this young man, right? How old is the baby?

MR. LYZENGA: Twelve weeks.

CO-CHAIR SEPTIMUS: Twelve weeks old. And he's still awake.

MR. LYZENGA: Thanks, Ed. A little tired, but, you know, hanging in there. Yes, go for it, Lisa.

MEMBER MCGIFFERT: Are these slides somewhere where we can download them? Are they on the page and then I can't see to find the big document you sent yesterday. I don't know who it came from, but I'm not -- yes, but who did the email come from because it's not from --

MR. ANDERSON: Yes, it came from me.

MEMBER MCGIFFERT: From you? Okay.

MR. ANDERSON: Yes, Drew.

MEMBER MCGIFFERT: It's not showing up.

MR. ANDERSON: And then it's also posted under the general documents section on the

committee SharePoint.

MEMBER MCGIFFERT: So what is it called?

MR. ANDERSON: It's called the Later than Competing.

MEMBER MCGIFFERT: The later --

MR. ANDERSON: Oh, the worksheets.

MEMBER MCGIFFERT: Oh, this is a later than --

MR. ANDERSON: It was an attachment.

MEMBER MCGIFFERT: It didn't --

MR. ANDERSON: Yes, I didn't. It's --

MS. QUINNONEZ: So --

MR. ANDERSON: But you --

MS. QUINNONEZ: -- if you go to the home page, it repeats.

MEMBER MCGIFFERT: Yes, the one that you -- the big document that has everybody's --

MR. ANDERSON: Oh, okay.

MEMBER MCGIFFERT: -- comments on it.

MR. ANDERSON: Yes, I can go ahead and upload it. I just sent it out as an attachment.

But I didn't --

MEMBER MCGIFFERT: As an attachment from you.

MR. ANDERSON: Yes, look back at the email that was sent late.

MEMBER MCGIFFERT: Okay. So I just did a search for you and I'll do it again.

CO-CHAIR SEPTIMUS: Look back at the email late yesterday afternoon --

MR. ANDERSON: Oh, that one came from Patient Safety.

CO-CHAIR SEPTIMUS: -- and it's got two attachments in it.

MEMBER MCGIFFERT: Okay. It's just --

CO-CHAIR SEPTIMUS: Did you not get it, Lisa? I can send it to you real quick if you don't have it. It should be late yesterday afternoon.

MEMBER MCGIFFERT: Well, I'll check it now.

CO-CHAIR SEPTIMUS: Isn't that right, Andrew?

MR. ANDERSON: Yes, it was.

CO-CHAIR SEPTIMUS: Yes.

MR. ANDERSON: It was around 4 yesterday.

MEMBER MCGIFFERT: Yes, it's -- great.

MR. LYZENGA: So --

MEMBER MCGIFFERT: Maybe you could see if you could send it to us maybe from your email?

CO-CHAIR SEPTIMUS: I'm going to send it to you, Lisa.

MEMBER MCGIFFERT: That's not --

MR. LYZENGA: Okay. So I was just going to take a quick moment to give you a refresher on our portfolio. You guys are again a pretty experienced committee. I think you're pretty familiar with our portfolio, so I won't spend too much time on this.

Just to note that we have had a little bit of attrition in the portfolio, some measures that have had endorsement removed, a few measures that have been withdrawn from consideration.

One notable area is the VTE area. We used to have a set of measures from the joint

commission around VTE, prophylaxis. They've decided that those are topped out, really getting high performance in those. So they've elected not to resubmit those for endorsement. So we've kind of had some shrinkage in that particular area, in a number of other areas as well.

Actually, if you look in the bottom corner there, this is a portfolio that actually is pretty heavy on outcomes compared to some of our others. It's not to say that we don't need more, but interesting to see that many of the other topic areas, portfolios of measures that NQF are not quite so heavy on the outcomes so.

I've got some slides here that walk through each of the topic areas and I'll probably just kind of skip over that in the interest of time.

We can get into our actual measure evaluation, but we can maybe return to this tomorrow when we get into our gap discussion, if you'd like, so we can look through how all of the measures that are actually in each of these particular topic areas. I won't belabor it at the

moment.

We've got some newly submitted measures. Most of the measures in this cycle of review are newly submitted, a bunch in that medication safety category, a couple of pressure ulcer measures, a couple fall measures, some HAI-related measures. Actually one of those is mis-categorized I think there, but. And then a few other sort of general or miscellaneous measures here. Do you have a comment, Ed?

CO-CHAIR SEPTIMUS: No, but I think if you go back and look at our portfolio --

MR. LYZENGA: Yes.

CO-CHAIR SEPTIMUS: -- I mean, look at the work that we've done over the last two/three years. It's a very impressive list of topics. So pat yourselves on the back.

MR. LYZENGA: So, yes, we can jump into the evaluation portion now, but before we get there, I did want to remind you we talked a little bit about this in our Q and A call, in our orientation call.

But we do have a new maintenance process in place wherein for maintenance measures we can, if the committee elects to do so, sort of skip over a couple of these or give a little bit less emphasis to a couple of these criteria. In particular, the sub-criteria under importance to measure and report of evidence, and then the scientific acceptability criterion.

And the idea is that once a measure has been endorsed, and for many of these measures has been endorsed multiple times, the evidence is unlikely to change a lot and nor is the testing. We've sort of already given that our okay through a series of committee reviews.

So we are allowing our committees to kind of skip over that portion without a vote. You may discuss it if you'd like and you may also vote if you decide that that would be appropriate.

In some instances there has been some, you know -- sometimes the evidence changes for something and in that case we would want to vote on evidence again.

Sometimes there's been updates in testing or updates to evidence, but often those updates to evidence and testing really only, sort of, add to the support for the measure or serve to strengthen the measure.

So in those cases even though if there's new information we don't necessarily need to vote again on those criteria. So we've done this in different ways across different committees.

We've only got two maintenance measures here, so maybe when we get to those we can just sort of do an informal hand vote on whether you want to maybe by exception, if you do want to vote on evidence for those measures maybe raise your hand. And if you do want to vote on scientific acceptability, raise your hand, but otherwise we can kind of pass over those criteria without much discussion.

CO-CHAIR THRAEN: So does the documentation have to demonstrate that there continues to be a gap in performance? I mean how is that evaluated in like with what you just said?

MR. LYZENGA: Yes, I should've clarified that. It's not in the entire importance to measure and report criteria just the evidence criteria and the sub criterion. We actually want to put more emphasis on the gap and the opportunity for improvement.

That is something we want to definitely do want to talk about and take a vote on for the maintenance measures and actually place a little bit more emphasis on that area. Any question on that from the committee?

MEMBER WEBB: I just need a voting device.

CO-CHAIR SEPTIMUS: You're very important. Any other questions?

MR. LYZENGA: Jason, you came in a little late. Did you get a voting device?

CO-CHAIR SEPTIMUS: Yes, Jason, will you, yes, introduce yourself --

MR. LYZENGA: Yes.

CO-CHAIR SEPTIMUS: -- and state any conflicts please, Jason?

MEMBER ADELMAN: Yes, I'm Jason Adelman. I'm the chief patient safety officer at Columbia University Medical Center and I don't have any questions. Oh, and I have no conflicts either.

CO-CHAIR SEPTIMUS: Is there anybody else who joined on the phone?

CO-CHAIR THRAEN: Do you have a voting device?

MEMBER ADELMAN: Yes, I'll vote.

CO-CHAIR SEPTIMUS: Did anybody else join on the phone that's a committee member? Okay. So I hope we do this all day, but we are, I almost hate to say it, but we're 20 minutes ahead of schedule.

But the first section this morning up until lunch Iona will be the moderator. So take it away, Iona.

CO-CHAIR THRAEN: Hey. So we're going to start with the consideration of candidate measures, 0022, use of high risk medications in the elderly, National Committee in Quality Assurance. Do we have a presenter from NCQA today for that?

And then Michelle -- where's Michelle?

MEMBER SCHREIBER: I'm here.

CO-CHAIR THRAEN: You are the presenter from the team after the NCQA presents. Thank you.

CO-CHAIR SEPTIMUS: Those chairs are wired. No, I'm kidding.

DR. GIAVONETTI: So, hi, my name is Erin Giavonetti. I'm a senior research scientist with the National Committee for Quality Assurance. I'm joined by Bob Rehm, an AVP in our performance measurement department. And on the phone we have Emily Morden who is the measure lead for this measure. Emily, you want to say hi?

MS. MORDEN: Hello. My name's Emily Morden. I'm a senior research associate with our performance measurement department as well.

DR. GIAVONETTI: So this measure is Use of High-Risk Medications in elderly is a maintenance measure. It is a long-standing HEDIS measure that was recently updated to match the updated American Geriatric Society Beers Criteria.

So the measure assesses whether or not older adults were dispensed a high-risk medication. There is extensive evidence showing that certain medications in older adults can be very harmful. They can result in adverse drug offence, falls, confusion, hospitalization, and even death.

The American Geriatric Society convened a panel of experts in geriatrics and pharmacology to review the evidence for medications which are harmful in the elderly and to create the Beers criteria which is a list of medications to be avoided.

The National Committee for Quality Assurance and the Pharmacy Quality Alliance, MCMS, were ex-officio members of that panel, non-voting members, so we were able to actually listen to the discussion of the evidence guideline developers.

We then took the Beers criteria and adapted it for use in a performance measure. The one that I have presented to you today is used in Medicare Advantage plans. It is completely

aligned with a parallel measure that is owned by the Pharmacy Quality Alliance that is used in Medicare Part D.

The measure has recently been updated by NCQA and went through review by our geriatric measurement advisory panel and our committee on performance measurement and was voted on by the board of directors at NCQA.

The updates we made to the measure, which you have in your materials is updated medication lists as well as an update to Rate 2 of the measure, which changed Rate 2 from looking at dispensing of two different high risk medications to be dispensing of two dispensing events for the same high risk medication. And this brings the measure in better alignment with the Pharmacy Quality Alliance measure.

And with that I will pass it back to the chairs.

CO-CHAIR THRAEN: Michelle?

MEMBER SCHREIBER: Okay. Well, thank you. And first of all I'm sorry I can't be with

all of you today. I've looked forward to being there, but I also had a family emergency, so thank you for allowing me to do this by phone.

I think we heard from NCQA that this is, indeed, a maintenance measure assessing the use of high risk medications in the elderly by the Beers criteria that has been in use really, for quite some time.

So I want to talk a little bit about the importance to measure first. And that's, you know, none of us actually would disagree that medication errors are really important and are among the top patient safety issues.

This measure assesses whether or not patients who are greater than age 65 and who are not in hospice, that's the one exclusion criteria, have been prescribed one or more potentially inappropriate medications.

The list is a very well-developed and well-referenced list called the Beers criteria. And as you heard they're based on recommendations from the American Geriatric Society, NCQA, CMS sat

in, pharmacology, and the medications that are associated really with numerous clinical trials and publications.

I believe the importance to measure is high because it represents potentially harmful medications to the elderly. But I do want to take a moment to comment on the public comment from CDC that came later with this measure.

Who, although they agree that medications errors are an important safety issue, they describe that this particular measure may not be best in identifying medications that lead to an adverse drug event in the elderly.

And their comment was that the majority of adverse drug events occur from warfarin, anti-diabetics or oral anti-platelets. My comment would be that I believe the importance for this measure is still high.

Warfarin is something that we prescribe because frankly you have to. And it's always given for a specific reason and it requires monitoring and that's the safety, I think, of warfarin.

This list does include key anti-diabetics such as glyburide, many of the CNS depressants such as barbiturates, sleeping medications and others. So they're included appropriately in this list and I believe the list still represents an important high-risk medication list and it's important to measure.

I'd also note that this harmonizes nicely with our next measure to be evaluated, 2993, which looks at specific diseases or risks. This measure is actually broader.

As we discuss this though I would maybe ask the committee that we not look as closely or specifically vote on reliability or validity because that has been tested over and over again and I think that's an opportunity for us to bypass that.

CO-CHAIR THRAEN: Ed has a comment.

CO-CHAIR SEPTIMUS: No, I was just clarifying. I'm sorry.

CO-CHAIR THRAEN: Thank you, Michelle.

MEMBER SCHREIBER: So --

CO-CHAIR THRAEN: NCQA, do you want to comment on the 0022 versus 2993 and clarify the differences between those two measures?

DR. GIAVONETTI: Sure. So in the Beers criteria there are multiple tables of medication. There is one table of medication, I know it's Table 2, that is the list of medications to be avoided regardless of condition.

There is an additional table that says if you have a specific condition these medications should be avoided because they either exacerbate the condition or can cause other problems. So an example there would be SSRIs and falls.

There's evidence that use of SSRIs leads to increased falls in the elderly. And so if you have a history of falls that should potentially be avoided. You would not necessarily say that would be something to be avoided for the whole population. So the measures are complementary, but they look at different aspects.

CO-CHAIR THRAEN: And then Steve.

MEMBER LAWLESS: A question actually

on Table 1A and you can interpret it for me. When you talk about the correlations and you said Rate 1, one high risk medication and Rate 2, two high risk medications, I would think that if you had one high risk medication you'd have a correlate.

And the correlation would be stronger if you had two and then sometimes the correlation gets, no, it's probably not significant, but it doesn't really change. So does this impact at all why one versus two?

DR. GIAVONETTI: So this is actually looking at one high risk medication and this is based off of the data we had available at the time that looked at two different high risk medications.

MEMBER LAWLESS: Got you.

DR. GIAVONETTI: So you are correct. They should be highly correlated. It was mostly a check on the validity. But if we did not see them to be highly correlated, that would suggest that there was a problem with the measure.

MEMBER LAWLESS: But I mean, it's almost like a tolerance thing has developed in

terms of one is more, but I don't see the -- even if you had two separate ones --

DR. GIAVONETTI: Uh-huh.

MEMBER LAWLESS: -- would you expect still to be more of a -- I mean, they're adding on to each other.

DR. GIAVONETTI: Yes, but this is looking at the health plan level rates, so you would expect to see that the rate of people who had one, right, that's going to be included then in the rate of people that had two. So they're overlapping measures. But the correlation therefore, they're not entirely independent rates.

So you are correct that a correlation is perhaps a little bit, you know, kind of fuzzy there which is why we've included the correlation with the other measures that use of high risk medications and specific conditions which are different high risk medications.

CO-CHAIR THRAEN: Would you also comment about performance gaps or performance? You said this is a maintenance stage. Has there

been much change over the last several years and if so, what direction?

DR. GIAVONETTI: So there has been change. It is a little bit challenging to look at a trend because the medication list is being updated. So there was actually a very big update to the measure in, I believe, 2012 where a good portion of the medications changed and we saw a big change in the rates. This time we've had another medication update, so it's hard, we'll not actually compare directly the two.

One thing we have noticed is that the second rate, which looked at the use of two different high-risk medications, we saw that decrease dramatically.

And we, therefore, felt that that measure was bottomed out in terms that there couldn't be much lower that they could go. And that's why we've revised it to look at the prescription of two dispensing events for the same high risk medication.

And that is aligning with the Pharmacy

Quality Alliance, but it's also because you may get somebody where they get the first prescription and that maybe was not completely avoidable, but when you are then giving them a second prescription for the same high-risk medication, particularly for things like sleeping agents, you are starting them down a pathway that can be very dangerous.

MEMBER SCHREIBER: It's Michelle. I'd also like to comment. I was going to talk about this when we voted about the performance gap, that there's actually a very nice table included in the measures that does show that there has been improvement over time.

So from 2012 to 2014, prescribing at least one high risk medication has fallen actually from a mean of 21 to 13.2. But if you look at the worst performers and the best performers, so the tenth percentile, in other words the top ten percentile, in terms of performance, that currently sits at 7.6, but the bottom tenth percentile is 21.7.

So I think there's still an opportunity

and a gap between the best performers and the worst performers in this measure.

CO-CHAIR THRAEN: Thanks, Michelle. So, Albert, then Charlotte.

MEMBER WU: Yes, hi. I just had, sort of, a little detail question about the Table 2, which actually, I think, here is shown for us as Table 1C.16.

There's a note that only prescription medications are to be included in the list. But in that table the first box includes anticholinergics and a lot of them are not prescription medications. So should those medications be included or not included?

DR. GIAVONETTI: Emily, you're closer to the actual individual medications. Can you clarify what out of that list? Are all of those included or is this just from the Beers, everything they listed as anticholinergics?

MS. MORDEN: So if you're referencing the table of medications that we included in the evidence for, these are all the medications we do

include in the measure. So the anticholinergics, for example, these first generation antihistamines, these are included in the measure.

And you're right that some of them are available over the counter, but some of them are also provided via prescription and so they are included. So we wouldn't include a medication, I guess, that wouldn't be dispensed via a prescription ever, if that makes sense.

CO-CHAIR THRAEN: So I guess the question is in the counting are you including non-prescribed medications because you have them listed or you're not including them because they're not prescribed? I think that's the question Albert's asking.

MEMBER WU: And to go a little further, since in electronic records medications are often listed even if they're not medications at all, are those then going to count against you or not, and are you going to have to discriminate whether or not they were prescribed or recommended or someone just said go to the pharmacy and get some of this,

it's cheaper because it's over the counter?

DR. GIAVONETTI: So let me clarify that this is not about prescriptions, this is about dispensing events. So if a medication was prescribed, but never filled, it would not be in the measure. If a medication was bought over the counter it would not be in the measure because it is based off of pharmacy claims data. Right now that is the data that most health plans have reliable access to.

I think it's an excellent point about the future of this measure and basing this measure off of data that is available in electronic health records that might include medications that somebody is on that were not prescribed, including the OTCs. And that will be certainly something we will investigate as we think about moving this measure towards using data from the EHR.

I think the other caveat to that is that we also know that there are sometimes medications missing from the medication list and the EHR that may be captured in the pharmacy data.

And we certainly would not want to create the unintended consequence of providers not listing a medication they're providing because they would be dinged on a quality measure.

So for example, I'm going to give you some antipsychotics, but I don't want to be dinged on the quality measures so I won't list it in the EHR. That would be the worst possible outcome, so.

MEMBER WU: There is a funny bias potentially that some people's insurance plans that are very generous will cover, free, a prescription of Benadryl, for example. And so people will ask to have a prescription written for a medication which might otherwise not be covered and, therefore, be bought over the counter.

So some people may be dinged for this and others not and that is unlikely to be in a random way.

CO-CHAIR THRAEN: Charlotte, then Yanling, then Ed, then Lisa.

MEMBER ALEXANDER: Thank you. Just a comment, I noticed in your Table 2, which is also

our 1C.16, that on a number of the medications the evidence is marked as low, but the recommendation was strong. And that seems a dichotomy to me, so could you speak to that, please?

DR. GIAVONETTI: So I'm going to try to -- they used the IOM criteria for grading evidence. And there is criteria that I think someone who's more of an expert in this could probably speak to it, but I believe, in order to rate something high or moderate there had to be a certain amount of randomized clinical trials or a certain type of trial that the evidence was based off of.

And there are some medications where there just isn't that evidence, but what the evidence that there is showed that the medication could have such a negative effect that there would not be any additional studies of that medication.

And so that's where the AGS came on. Even though the evidence was low, you know, in terms of the design of the studies, the actual impact of the medication was so high that they felt it was a strong recommendation.

MEMBER YU: Thank you. My question's for the developer. I think this is very important measure and definitely that shows gaps to improve, and also it shows the improvement from 2012 to 2014.

My question to you, there are two questions, first one, in your analysis what is the most critical if you could list one or two factors that contributed to this improvement during this time window? That's the first question.

DR. GIAVONETTI: So I think the most critical factor was the inclusion of PQAs parallel measure of use of high risk medication in the elderly in the CMS Part D stars rating program.

And that is a program whereby plans received financial benefit for improvement on quality measures. And as we saw this measure implemented in that, we saw the rates go down significantly. So I think that's why we are seeing to purely this measure going down.

Did you want to know what, specifically, our health plan's doing to reduce the use of this medication?

MEMBER YU: That actually is my second question.

DR. GIAVONETTI: Yes. Uh-huh.

MEMBER YU: What is the adaptation rate for, you know, the implementation and adaptation for this measure for, you know, CMS, you know, that we all know they have lots of -- I'm more interesting about the commercial, the other type of health care plan and also in term of traditionally in for physician, was it called the physician reporting system?

DR. GIAVONETTI: Uh-huh.

MEMBER YU: So I'm interesting to know particularly on that.

DR. GIAVONETTI: So this is a measure that is reported by all Medicare Advantage plans that have a pharmacy benefit to them. It is required by CMS that they all report on that.

In terms of the adaptation by physicians, there is a version of this measure that is in the physician level version. It's not the one that we're bringing to you today.

I can't speak to all the details of it right now, but it is a measure that is being collected through EHRs and is slightly different, but I don't have the numbers in front of me in terms of how many providers are actually using that measure.

And there are a different set of issues associated with that measure, but what we're focused on today is the health plan level measure.

MR. REHM: And if I can add a little bit to that. The HES when it first came out with its revised guidance in 2011 really upped the ante by providing a whole variety of clinician tools and explanatory articles in addition to the guidance.

And I think they also, under our recommendation, had a public comment period which I think helped pollinate the guidance. And then when that was repeated in 2015 again, stepped up their communication strategy around this.

So this was an example of where the AGS went really out of its way to provide a variety of tools that both patients and providers can use.

And so if you're thinking about kind of a trickle down from health plan, which is a fairly large population-based measure to getting traction in the clinical community, I really do think that there's a trend here and I think stars is important. But I also think clinician engagement is important as well as patient understanding and appreciation.

MEMBER YU: So in other words, you anticipate there would be increased adaptation of this measure at the healthcare system level?

MR. REHM: I think it's kind of like, you know, do you usually run out and buy the first year model of a car or do you go for the second year. And as these things develop, I think the traction just builds. It builds on itself.

MEMBER YU: Yes, I was just --

MEMBER SCHREIBER: Actually, this is Michelle, I'm sorry because I can't see you to know who's commenting and talking.

But I would add that as an integrated care delivery system, which we are, although currently this affects our health plan, all of our

physicians have been included in this and in the prescribing parameters on what to avoid in the elderly. So I do think that from a personal practicing point of view that's true.

In addition, this is being used as a PQRS measure and as a meaningful use measure and so I think it will affect and certainly trickle down to individual providers.

MEMBER YU: Okay. Thank you.

CO-CHAIR THRAEN: Ed and then, Lisa, and then Kendall, right?

CO-CHAIR SEPTIMUS: Yanling actually asked almost all of my questions, so that's scary. But one question for my own information. Is 65 the right age?

CO-CHAIR THRAEN: For what?

CO-CHAIR SEPTIMUS: Now, part of this is personal guys.

MR. REHM: It's very personal to some of us.

CO-CHAIR SEPTIMUS: But, you know, 65 today is a little different than it used to be. And

my question is, is 65 the right cutoff for this? But all my other questions were asked by Yanling.

DR. GIAVONETTI: Thank you. You raise a really good point that I will bring back to the American Geriatric Society, who, you know, they're also experiencing the same aging.

Is it the right cut off? It is in some ways, 65 is an arbitrary cut off that is used commonly. And I think that's where we see the research focus on population 65 and older.

I think as we see this silver tsunami, you know, we're going to see, probably, more research saying 65 isn't the same as it used to be. And maybe there might be some, but we don't want to get ahead of the evidence.

CO-CHAIR SEPTIMUS: No.

DR. GIAVONETTI: Uh-huh. Yes.

CO-CHAIR SEPTIMUS: It's not the question's too big.

DR. GIAVONETTI: Yes.

MEMBER MCGIFFERT: It's also related to the Medicare age, so that probably has something

to do with it.

I had a couple of questions, and maybe you've already answered it, but I see that denied claims are excluded and I wonder if that is missing a lot of inappropriate prescriptions.

Maybe it's more appropriate to include denied claims when you're looking at it from the physician perspective, but also if you had in a plan it seems if you had a lot of physicians prescribing it and then it was denied that that's an issue to consider.

DR. GIAVONETTI: So I'm going to invite Bob and Emily to speak on this. I think one of the things that we are aware of is that one of the tools that health plans have to influence provider behavior is denial or asking for preauthorization for certain medications. That is one of the ways that they can push providers to say don't prescribe this. You need to justify why you're doing this.

That's why if a health plan is doing all of that and providers are still, you know, filling a prescription, that's where we kind of have said

at the health plan level, they have met the intent of not covering this medication.

At a provider level I do agree that a denied, you know, claim should not be excluded. But I would ask Bob or Emily if they want to add anything.

MR. REHM: No, I think you've captured it. It's the famous tools in your tool chest. What can you do? I think that just the idea that they can do has an influence over behaviors, not that they do do. And it's not that it's a common event.

And if you've seen a health plan and you've seen the appeals process for a variety of things, most appeals are overturned in the long run. They just don't really want to work it out. But we want to make sure that we give credit where credit's due.

But totally true from the optics of a clinician level measure, which this is not, we're not seeking endorsement at that level. Your point is well-taken. From a health plan perspective

it's the right thing to do.

MEMBER McGIFFERT: Many years ago consumers union, we did a little study in Texas on the appeals process there after it was in place for a few years. And it was interesting that the kind of denials that were not overturned were on the prescription drugs, you know, that seemed when an independent person evaluated it, those were less frequently overturned.

I had another question about the disparities issues. And I may have missed the details, but it looked like you were saying that there is a way to determine disparities, but that you hadn't done it, or am I reading that wrong?

DR. GIAVONETTI: So, as a HEDIS measure, all HEDIS measures are reported at the health plan level and are not currently reported out by race or ethnicity.

This is certainly something that we are looking towards finding better ways to have these types of data reported in the future. It is dependent in a large way on health plans having

accurate data about race and ethnicity, which is a little bit better in Medicare than it is in commercial and Medicaid plans.

MEMBER WEBB: So my question was also along the lines of disparities, if you will, but not necessarily on race, just plain socioeconomic status. I see a lot of places in the chart where it says there are other alternatives available.

And was there anything done to, for instance, I work in a safety net hospital and there are a lot of times we have to give a patient Benadryl instead of hydroxyzine because they can't afford the hydroxyzine, although both of those are on the list.

But is there any method of determining, for instance, or risk stratifying somehow? You know, not all health plans are the same, so not all health plans are going to cover the alternative medication that you guys are talking about. And how is that considered as part of this measure?

DR. GIAVONETTI: So as a health plan level measure, we see this is something that is

within the health plan's control to cover the medications that are the alternatives or to make those medications more affordable to the patients so that the Benadryl is not the most affordable option.

So that's something where we actually see the health plan could play a role in trying to decrease disparities in the receipt of these medications by helping to make these options more affordable.

MEMBER WEBB: And how would this measure help the health plan understand that that's what they need to do?

DR. GIAVONETTI: So health plans often look at their own rates stratified by a million different things. They do their own investigation in why is my rate on this measure so high. And that is something we know health plans are doing.

It is not something that we get from HEDIS data because there are so many different issues which this committee is probably well-aware of, NQF is very well-aware of, of how do you

actually define lower income from administrative claims data that we have available. How do you not mis-classify people?

And, then, is that a potential that you're setting to different standards if you have two different rates, saying it's okay, you know, for this population to have more use of high-risk medication.

So this is continuing work that's going on. There's been lots of exploration into this, this measure included in that work. But what we're bringing to you today is the measure as it is used in HEDIS right now which is not reported by separately in this.

MEMBER QUIGLEY: Thank you, Madam Chair. Pat Quigley for those on the phone to know who's speaking. And, Michelle, thank you for your comments.

And my comments in relationship to this measure that has been around for some time as a process measure is really to help think about how this measure is going to go forward because there

is a difference between polymedicine and polypharmacy. And having one versus two medications is not necessarily an indicator of quality when you think about polymedicine versus polypharmacy.

In polypharmacy many medications -- an older person can be taking because they're going to different providers. But in polymedicine you have to have the right medication to treat the right comorbidity and then consider the interaction of those medications with those other medications to treat a patient. So I think that there is work to be done to really move this measure to quality.

And the other comment that I would like to make is in relationship to the population that has been excluded from this measure, and that is the hospice patient population.

That now, starts moving into a population-based approach to looking at quality measures. And there are other patient populations in the aging population that this measure is not relevant for.

If you think about the older patient who falls and has a spinal cord injury, spinal cord injury in old people, as a new diagnosis the number one cause is a fall. They are going to be on these medications, spinal cord injury patients.

If you think of the next population that I'd like to mention and I could continue to go on, is the aging mental health patient, geriatric psychiatry patients. So sometimes with these patients just to lower a dose or to change a medication and a class is a quality measure.

So I just want to make that comment publicly that this has been around for some time, but to just continue to still focus on one versus two medications really is more on the polypharmacy versus the polymedicine side. So those are the comments that I'd like to make. Thank you very much.

CO-CHAIR THRAEN: Any other comments or questions? I think we have to vote next, right? Is that the next process?

MR. LYZENGA: Yes, so we're starting to

veer a little bit into some of the specifications issues. So maybe we should just ask if do we want to vote on evidence as a committee or are we comfortable accepting the previous committee's acceptance of the evidence on this measure?

Is there anybody that does want to vote on evidence? All right. Seeing none, I believe we do want to vote on opportunity for improvement in gap and care. So we can do that and then we'll move on to the reliability section.

CO-CHAIR THRAEN: So I think we need a refresher course on how to use the gizmo.

MR. LYZENGA: Yes, good call.

MS. QUINNONEZ: Okay. I'll just provide you with a few instructions for your voting clicker, just a few reminders. So when it's time to vote, if you would pick up your clicker and point it directly towards me or to this laptop down this way.

Okay. Now, each individual click, each one of you, each clicker, will hold one vote. You can click it as many times as you'd like to,

but the computer will calculate and capture the final click that you do, okay?

And also, we may have to click multiple times, but just bear with us, it's a part of the fun process. Okay. All right. Let's see.

PARTICIPANT: Wait a minute. Which do we vote for, this one now?

MS. QUINNONEZ: No, we're not voting yet. I'll tell you when we're ready to vote. I'm sorry.

MEMBER SCHREIBER: And are you going to put on the screen what we're voting on? Because otherwise I can't see it.

MS. QUINNONEZ: Yes, Michelle. Sorry. I will read it out to you because we can't share that screen particularly. But I will read out what we'll be voting on, okay?

MEMBER SCHREIBER: Great. Thanks.

MS. QUINNONEZ: And you can actually submit your vote through the chat box.

MEMBER SCHREIBER: Yes, that's what I was going to do. Thank you.

MS. QUINNONEZ: Okay. Perfect. Okay. So we are now going to be voting on the importance to measure and report for Measure 0022. Voting criteria is 1 will be high, 2 will be moderate, 3 will be low, and 4 will be insufficient.

Voting is now open and you can vote on the importance to measure and report requirement gaps for Measure 0022. Now, we believe we're looking for 20 votes, so we need just a few more. They're not in order.

CO-CHAIR SEPTIMUS: Still don't have all?

MS. QUINNONEZ: We're looking for one more vote. And you did Michelle's?

CO-CHAIR SEPTIMUS: The light's got to light up for it.

MS. QUINNONEZ: Can you do roll one more time just to make sure? I don't have 20 in house right? 19 of those. Okay. All right. Voting is now closed. We have 63 percent voted high, 37 percent voted moderate, 0 percent for low and 0 percent for insufficient. All right.

MR. LYZENGA: Okay. So we can move on --

MS. QUINNONEZ: I think that's a composite.

CO-CHAIR THRAEN: No, we'll skip that slide.

MR. LYZENGA: Right. So we're going to move onto scientific acceptability. Michelle?

MEMBER SCHREIBER: Okay. I didn't know if we were going to skip this part or not.

MR. LYZENGA: Oh.

MEMBER SCHREIBER: Under scientific acceptability reliability the measure is currently used in HEDIS and a review from HEDIS from 2012 to 2014 was used to calculate reliability with the binomial method.

And it demonstrates reliability measure for one prescription at .99882 and for two or more is .99819, so reliability appears high. Do you want me to do validity at the same time?

MR. LYZENGA: Sure.

MEMBER SCHREIBER: Okay. So --

MR. LYZENGA: Oh, do you want them separately? Sorry, no. One at a time.

MEMBER SCHREIBER: Okay. Sorry.

MS. QUINNONEZ: Okay. We will now be opening voting for reliability of Measure 0022. Voting is now open.

PARTICIPANT: It's not working.

MS. QUINNONEZ: Okay. Okay. I see votes coming in.

MR. LYZENGA: So we actually have the option, I just informed, we have the option of deciding not to vote on this as well. Sorry.

MS. MUNTHALI: And just to clarify, the reason we're giving you that option is because testing hasn't changed since the measure was last reviewed. And as you remember, Drew and Andrew were telling you we have a new maintenance process. And you can opt as a committee on whether or not you want to vote on this criteria, so just a yes or no.

CO-CHAIR THRAEN: Okay. So does anybody want to review reliability and validity?

We'll start opposite? Okay. We'll move forward then.

MR. ANDERSON: Yes, so you can discuss feasibility.

MEMBER SCHREIBER: Okay. So we've moved past reliability and validity. Feasibility is also high. This is already being done. And as pointed out, it uses administrative claims data that is widely available. So are we then voting on feasibility?

CO-CHAIR THRAEN: Yes, we will.

MEMBER SCHREIBER: Okay.

MS. QUINNONEZ: Okay. Voting is now open for the feasibility of Measure 0022. Option number 1 is high, option number 2 is moderate, option number 3 is low, and option number 4 is insufficient. Okay. All votes are in and voting is now closed. We have 100 percent voted high.

MEMBER SCHREIBER: Do you want me to speak next to usability?

MR. ANDERSON: Yes, please.

MEMBER SCHREIBER: Okay. Sorry, it's

just hard not seeing all of you. So usability is also high. The measure's really already used in plans for NCQA, for health plan report cards, for the Medicare CMS star rating for health plans, for ACO accreditation, for PQRS, and for meaningful use.

MS. QUINNONEZ: Okay. Voting is now open for usability and use. Option number 1 is high, option number 2 is moderate, option number 3 low, and option number 4 insufficient information.

Looking for two more votes. All votes are in and voting is now closed. The vote reads 85 percent high, 15 percent moderate, 0 percent low and 0 percent for insufficient.

MR. LYZENGA: If I could just jump in, we sort of made a mistake. We can skip over the validity testing or the validity vote if you want to. There are some things that we could discuss around risk adjustment exclusions.

We talked about that a little bit already, SDS adjustment which is maybe not as

applicable for this as a process measure, arguably. But I just wanted to check if there's any of those issues that anybody wants to discuss and revisit the validity criteria. And if not, we can move on, but I just wanted to raise that as a possibility. Pat?

MEMBER QUIGLEY: Thank you. Pat Quigley. My comments were about increasing the exclusion criteria, was part of what my discussion was. So just that does include in that topic area. There's more than just the hospice patient part of this.

MR. LYZENGA: Is there anybody that wants to hold a vote on validity? I think we did hear the comments on the exclusions? No. Okay. Seeing none, we can go on to overall suitability.

MEMBER SCHREIBER: Overall suitability I believe is --

CO-CHAIR THRAEN: Michelle, just a minute. Just a minute.

MEMBER SCHREIBER: I'm sorry.

CO-CHAIR THRAEN: I just want to make

a note in the minutes that this won't come around for another three years. And we may not be the same people sitting at the table.

So I just wanted to make a note in the minutes that those issues that were identified, the demographic issues, the health disparities issues, and the patient population issues be addressed by in the next iteration of this process, in the minutes. Thank you.

MEMBER LAWLESS: You guys, just some more clarification for me on what are the exclusions in hospice, is there an exclusion with palliative care?

DR. GIAVONETTI: No, there is not an exclusion for palliative care. The reasons for this are twofold. One is that this measure is because we are trying to align with the pharmacy quality alliance measure which is for Part D plans, it only is based off of Medicare enrollment data and pharmacy claim data. We don't use medical claim data. So it's not feasible if we want the measures to remain aligned.

The other issue, and we have been discussing this with our expert panels, is palliative care is difficult to define as a population because there may be some people who are receiving some type of palliative care, but they would not necessarily be someone you'd want to exclude from this measure. There are different degrees.

It's an ongoing effort we actually have across all of our HEDIS measures to evaluate what types of advanced illnesses that are not hospice should be excluded from HEDIS measures.

It's really challenging because claims data just don't include that information. But I do hope that the next time this comes around we may be able to talk to you more about some exclusions for people that are clearly near the end of life, but may not be in hospice yet.

MR. REHM: And if I can just add, Pat, if you said you had a long list of conditions, like spinal cord injury, but similar ones, to the extent you want to share that with AGS, I'm sure they'd

appreciate it because I think that's perspective. You know, it's a small population, but that doesn't mean it's not an unimportant population, so.

CO-CHAIR THRAEN: Yanling?

MEMBER YU: Just quick comments to echo Pat about the exclusion of the hospice patient. I did make notes too. I'm a little concerned about the exclusion of hospice patients also as well because overall the goal is to improve the quality of care.

And for hospice patients they have their special need. The quality may be different, but the thing is still to this population patient is to reduce unnecessary complication and, you know, the harm also for this vulnerable population.

Some people can live on hospice for over years if you have a good quality care. Therefore, I think it's still an important thing to, sometime down the road, to think about it.

DR. GIAVONETTI: So the reason we have a hospice exclusion is twofold. One, is that the AGS in the Beers criteria actually stipulate that

this evidence does not apply to individuals that are in hospice or end of life.

So we didn't feel that we had the sufficient evidence to say that the risks of these medications in that population outweighed the benefits.

The other reason has to do with just the feasibility of individuals who are in hospice often are no longer under the control of the Medicare Advantage plan. They're receiving their hospice benefit through a Part A.

They may stay in the Medicare Advantage plan for supplemental benefits, but the plan is not responsible for their medication. So those are the two reasons why we exclude hospice.

CO-CHAIR THRAEN: Okay. Shall we vote?

MEMBER SCHREIBER: I'm sorry, can you reiterate what we're voting on this time?

MS. QUINNONEZ: Absolutely. Voting is now open for the overall suitability recommendation for endorsement for Measure 0022.

And Option Number 1 is yes, Option Number 2 is no. Okay. All votes are in and voting is now closed. The vote reads 100 percent for yes, 0 percent for no.

CO-CHAIR THRAEN: Albert.

MEMBER WU: This is just a comment for, you know, sort of thinking about ways if it's feasible for exclusions. I have one patient offhand who has intractable seizures. He's been on everything and he takes phenobarbital. And it's the thing that, you know, it keeps him from having seizures and falling down.

So, you know, he is someone who I'm going to continue to prescribe for and will get dinged for him, but it doesn't seem quite right. So just sort of thinking about ways to opt people out would be something to think on.

DR. GIAVONETTI: And that's also one reason why we never would want the rate on this to be zero. There will always be situations where these medications are appropriate. I think what we can say though is that 20 percent is not good.

CO-CHAIR SEPTIMUS: All right. No one? Okay. We're ahead of schedule and for those who want to take a bio-break this is a good time to do it, but you guys have great bladders. We'll keep going.

CO-CHAIR THRAEN: All right. So we have NQF at the table so we need to torture them a -- not NQF, NQCA, we need to torture them a little bit more for 2993, potential harmful drug disease interactions in the elderly. And Theresa is the lead on this. And you want to go ahead and do your presentation and then we'll turn it over to Theresa?

DR. GIAVONETTI: So I'll keep this pretty short. And most of the things we said in the previous measure apply to this one. The one difference here is that this a measure that is only reported by Medicare Advantage Part C plus D plans. So this is not a Part D measure.

Medicare Advantage Part C plus D, so it only includes people who have a pharmacy benefit and a medical benefit from the health plan. So if

you have a --

PARTICIPANT: If you have a managed plan.

DR. GIAVONETTI: Yes, in managed care.

CO-CHAIR SEPTIMUS: So can I have the mic?

DR. GIAVONETTI: Yes.

CO-CHAIR SEPTIMUS: One more on education, are there Medicare Advantage plans that don't have that C and D also, just for my own information.

DR. GIAVONETTI: There are a few, yes.

CO-CHAIR SEPTIMUS: Okay.

MR. REHM: But this excludes just Part D only.

DR. GIAVONETTI: So this measure is based off of the same evidence the Beers criteria. It looks specifically at there's four rates.

It looks for the people who have a history of falls and received a high risk medication, those with dementia and received a potentially harmful medication, and those with

chronic kidney disease and received medication. And then there is a total rate that combines all three rates.

This measure, in particular, the one thing I would like to highlight about this is that I think the performance rates for this show a real gap in performance and a need for improvement.

We see particularly high rates of inappropriate medication for those with a history of falls and dementia, which I'm trying to get to, with 48 percent of people with a history of falls getting a potentially inappropriate medication and 48 percent of those with dementia, so. Sure.

CO-CHAIR THRAEN: Theresa.

MEMBER EDELSTEIN: Okay. I don't want to repeat everything she just said in the first section, so as you know this is a process measure. The evidence is the same as the measure we just discussed. So if there are questions about the evidence that we haven't covered we should talk about those.

As was noted, the opportunity for

improvement is significant. It is more compelling for people with history of falls and fracture as well as people with cognitive impairment or dementia, less so for those with chronic kidney disease.

Do you want to stop there or should I go forward?

CO-CHAIR THRAEN: Albert, you have your sign up. Did you have a question?

MEMBER WU: I have a questions about sort of temporal relationship. So I have someone who falls down, who sustains a spinal cord injury and is put on muscle relaxants. Do I get dinged for that patient?

DR. GIAVONETTI: Emily, please correct me if I'm wrong, but if there is a fall and it's documented in the claims for accidental fall, we do not have an exclusion for spinal cord injury in that particular measure. Emily, am I correct in that?

MEMBER WU: And I don't just -- spinal cord injury, I mean has a seizure, has something

which then might lead to a potentially legitimate prescription of one of the medicines on the Beers list.

DR. GIAVONETTI: I believe seizures is an exclusion from that particular rate because anticonvulsants are on that and that would be an appropriate medication where the risks of not being on anticonvulsant if you have seizures would outweigh the risks of the falls.

But no, there is no analysis of the temporal sequence of things. It looks for if in the measurement year you had this condition and you received this medication.

CO-CHAIR SEPTIMUS: Quick question. Is this not a new measure?

DR. GIAVONETTI: This is new measure to NQF. This is a long standing measure in HEDIS, yes.

CO-CHAIR SEPTIMUS: No, but, you know, this is new to NQF. I just wanted to make --

DR. GIAVONETTI: Yes, this is new to NQF.

CO-CHAIR SEPTIMUS: -- that distinction.

CO-CHAIR THRAEN: Lisa.

MEMBER McGIFFERT: How do you define history of falls?

DR. GIAVONETTI: Wow. I so wish we could define it better. This measure is based off of administrative claims. So we look for a claim for an accidental fall or a hip fracture because as a proxy that most hip fractures are the result of a fall.

We know that we're probably undercounting falls because people show up and they have a history of falls and it's not going to show up in an ICD-10 or ICD-9 code. But one thing we do think is that those people with falls that result in injuries are more likely to be captured in this measure than those people who have a history of falls that maybe not resulted in injuries.

So it's not perfect. It is what we think is the best we can do with the administrative claims data that we are using.

CO-CHAIR THRAEN: Yanling.

MEMBER YU: Thank you. First of all it's a technical question. It's on Page 4. Maybe I just don't understand the stated part is the reliability testing and it shows a rate of Rate 1, Rate 2, Rate 3, Rate 4. Rate 4 is total. And then you have beta-binomial rates. And the rate for the total, the rate is 98.9857 which is higher than any individual ones. So --

CO-CHAIR SEPTIMUS: Sorry, I don't want to cut you off. We probably want to cover evidence and opportunity for improvement first and then get into --

MEMBER YU: Oh, I'm sorry.

CO-CHAIR SEPTIMUS: -- the reliability thing.

MEMBER YU: That's a totally different --

CO-CHAIR SEPTIMUS: Yes.

MEMBER YU: Okay. I'm sorry.

CO-CHAIR SEPTIMUS: No problem.

MEMBER YU: I jumped my -- I take back

the question.

CO-CHAIR THRAEN: Steve?

CO-CHAIR SEPTIMUS: Well, we can --

MEMBER LAWLESS: Does everybody, because it sounds like we're confused here, do we need a quick review how we review this or is everybody -- so can we just -- is it okay? Just let's find because it sounds like we've sort of forgotten --

MR. LYZENGA: Yes. Yes.

MEMBER LAWLESS: -- which is understandable.

MR. LYZENGA: That's okay. Yes.

MEMBER LAWLESS: If we did that it may help the discussion later.

MR. LYZENGA: Yes. Apologies for not doing that before. We try to, as much as possible, walk through each of the criteria sequentially.

So we'll want to talk about evidence first, have a discussion about evidence, vote on the evidence, talk about opportunity for improvement, take a vote on it and then move to

reliability, take a vote and then so no throughout the criteria until we get to the overall vote.

CO-CHAIR THRAEN: Steve.

MEMBER LAWLESS: Yes, question, this is on where the gap was there a problem. It may be the wording. I says 48 percent of individuals with a history of falls have high risk. What's the rate of falls in people with none of these conditions? Does the rate of falls of 52 percent have a history of falls who have none of this? So there's almost like an equal rate?

DR. GIAVONETTI: Right. So that half of people with a history of falls are receiving one of these medications.

MEMBER LAWLESS: And half are not?

DR. GIAVONETTI: And half are not.

MEMBER LAWLESS: And what is the rate of falls in people who are not -- have any of these conditions?

DR. GIAVONETTI: So you mean what is the rate of people with a history of falls in the plan?

MEMBER LAWLESS: Right. With none of these conditions or just almost like is there a gap? So if somebody does not have dementia or kidney disease, in this population what percent of those patients have falls?

MR. REHM: So falls is the condition we're looking at, if you will. Dementia is a separate condition. Kidney disease is another condition.

So there's a population of people who have not fallen, they're not in the measure. And there's a population of people that have fallen that are in the measure on this indicator. So it's -- we're not combining falls with anything else right now. Does that help?

MEMBER LAWLESS: I'm trying to look at the gap and so you qualified it.

MR. REHM: Okay.

MEMBER LAWLESS: So if falls is the thing, going back to the question that was asked, the interval of the fall documentation versus the medication documentation.

MR. REHM: Right.

DR. GIAVONETTI: So we look for --

MS. MORDEN: Erin, I can clarify that if it would be helpful.

DR. GIAVONETTI: Thank you, Emily.

MS. MORDEN: So for the fall we are looking for anyone that had a fall between basically, January 1st of the prior year up through December 1st of the measurement year. So we have a window there where we're looking for did a fall occur.

Then for the potentially harmful medication, we're looking to see if that was dispensed after the date of that fall up through the end of the measurement year. So we are looking for the dispensing of that medication to occur after the fall.

CO-CHAIR THRAEN: Steve's shaking his head.

MEMBER LAWLESS: So therefore, but if the medication was prescribed after the fall, this measure kind of implies that the medication caused

the fall.

DR. GIAVONETTI: No. So just to clarify, this medication, this is based off of the evidence that shows people with a history of falls should not be prescribed one of these medications.

Now, one of the reasons they should not be prescribed this medication is because it would cause more falls. So that's where we start with people with a fall and look to see that they get one of these medications.

CO-CHAIR THRAEN: Pat and then, Yanling.

MEMBER QUIGLEY: Thank you. Pat Quigley. And my comments are related to this really being grounded in the AGS guidelines again.

And it's in the algorithm of the AGS guidelines, and we discussed this at the last time we had talked about a community-based fall risk assessment measure, is that the guidelines ask if a patient's had a fall in the last year.

But the patient's who actually get worked up for evaluation are patient's who have had

more than one fall in the last year or an injurious fall. And that's the more vulnerable side of the AGS guidelines.

And last year the United States Preventative Services Task Force recommended that those that do get worked up are those that have had more than one fall or an injurious fall, not just one fall in the last year because there was no evidence to support the burden of evaluating patients.

So I would just say certainly this is an important indicator, but the measure would have had more opportunity to improve quality if it was focused on those who are more vulnerable, those who are falling more than once or had an injurious fall.

And that hip fractures are not the only injury. It's head injuries. Older people who fall and have head injuries is just as debilitating if not, in terms of mortality and morbidity, as hip fractures. So that would be the comment I would make in relationship to the target population.

CO-CHAIR THRAEN: Yanling.

MEMBER YU: Thank you. Now I'm back on track. The question is, you know, there's a condition on the list as dementia diagnosis. So I was just wondering, you know, from what I know there are lots of elderly, especially in a long-term facility, they are on the boundary of, you know, dementia or just because aging confused the health condition.

They're described a loss with those improper drugs, like a psych drugs. So how do you think this would be captured as a -- I think it's quite common in the facilities, especially long term. How do we capture this type of thing, rather than just a diagnosis of, you know, dementia?

CO-CHAIR SEPTIMUS: Turn yours up. You're good. You're good. It just came up red.

DR. GIAVONETTI: Oh great. You're correct. We are not capturing the people that there is no diagnosis and claims for dementia. And then we know that dementia is under-reported in claims or under-diagnosed.

I think as we look towards the future

of this measure that could be based on either claims data or EHR data where we may have a diagnosis in a list of problems in an EHR that includes dementia, that that's a possibility for this measure.

Right now, we don't have as clear evidence from the American Geriatric Society that pre-dementia, mild cognitive impairment is --- there's evidence of the risks of these medications. So they focus strictly on those people with a dementia diagnosis in their evidence review.

We may start to see more evidence come out on the risks of these medications and those people before they have a diagnosis of dementia. Does that answer your question?

MEMBER YU: Yes.

CO-CHAIR THRAEN: Kimberly.

MEMBER APPLEGATE: Just a quick reminder that we reviewed another measure last year that addressed some of the concern about psychometric medication and restraints. And I know it doesn't exactly address this, but in long-term care facilities.

So it gets at some of this issue around inappropriate use of psychiatric medications for sedation of patients. So it may not address this metric, but it does --- there's another metric out there that is getting at this issue. Thanks.

CO-CHAIR THRAEN: Any other questions before we call for the vote? All right. Should we vote on evidence?

MS. QUINNONEZ: Voting is now open for Measure 2993. And we're voting on the importance to measure and report. Voting Option Number 1 is yes, voting Option Number 2 is no. We should be on the one that --

MR. LYZENGA: We should be on the process --

MS. QUINNONEZ: Got it.

MR. LYZENGA: -- measure, which has a few different options actually.

MS. QUINNONEZ: Yes, sorry. Okay. Give me one second. Hold on here. I'm going to get -- okay. We're going to vote now on the importance to measure and report for evidence

structure, process, and immediate and outcome measures.

So we're voting on the evidence of Measure 2993. Voting Option Number 1 is high, voting Option Number 2 is moderate, voting Option Number 3 is low, and voting Option Number 4 is insufficient. You may place your votes.

Okay. All votes are in and voting is now closed. Vote for evidence of Measure 2993, it reads 65 percent voted high, 35 percent voted moderate, 0 for low percentage, and 0 percent for insufficient.

CO-CHAIR THRAEN: All right. Theresa, you want to cover validity and reliability or is -- this is a new measure, correct?

MR. LYZENGA: Yes.

CO-CHAIR THRAEN: So we need to go through that.

MR. LYZENGA: Yes, performance gaps.

CO-CHAIR THRAEN: I'm sorry, what? Performance gaps, sorry.

MEMBER EDELSTEIN: Okay. Performance

gap it is. Okay. So as we already spoke of, in the 2014 data was shown that there is a sizable gap between health plans at the 10th percentile versus the 90th percentile, so the opportunity for improvement appears to be high.

CO-CHAIR THRAEN: Are there any questions on performance gap? Shall we vote?

MS. QUINNONEZ: Voting is now open for the importance to measure and report performance gaps for Measure 2993. Voting Option Number 1 is high, voting Option Number 2 is moderate, Option Number 3 is low, and voting Option Number 4 is insufficient.

We're looking for two more votes. All votes are in and voting is now closed. The vote for performance gaps of Measure 2993 reads 85 percent voted high, 15 percent voted moderate, 0 for low and 0 for insufficient.

CO-CHAIR THRAEN: So, reliability.

MEMBER EDELSTEIN: Okay. I just want to clarify on the sheet there's a section on priority, is there a separate vote for priority?

No, okay. Making sure.

Okay. So under reliability, the measure --- reliability was done at the measure score level using a beta-binomial testing. They used 2014 health plan data that covered 412 health plans. The level of reliability based on the results appears to be high.

CO-CHAIR THRAEN: Anybody have any questions or comments? Yanling, go ahead. You had a technical question earlier, you want to --

MEMBER YU: Yes, that's my question about, so I don't need it.

DR. GIAVONETTI: So, the reason that the reliability for the total rate is higher is because the denominator for that rate is larger than the denominators for the other rates. So you've got more population and that influences the rate of the calculation of reliability --

MEMBER YU: Okay.

DR. GIAVONETTI: -- to the mix. Does that answer your question?

MEMBER YU: Yes, that's fine. Thank

you.

CO-CHAIR THRAEN: Albert?

MEMBER WU: And as far -- I heard of reliability of lots of different components. What is the reliability for the ascertainment of a fall?

DR. GIAVONETTI: So we don't have an assessment of that. Our reliability is at the score level not at the individual item level. So we do not go back and, I can't recall if we did in the long ago metric testing of this, look at the claims of falls versus falls in the medical records. But for right now, we're in the --

MEMBER WU: So to validity then, did we --- were there tests of the validity of the ascertainment of a fall?

DR. GIAVONETTI: No, the validity is also at the measure performance level. So we looked at construct validity of the performance of the measure compared to other measures.

MEMBER WU: Are you satisfied with that?

DR. GIAVONETTI: Well, having been

taught by one of the experts in measure testing on testing of validity and reliability, I will say that I don't think -- no names. Do I think that the claims for falls is the best measure of falls, no. But, it is what we have available to us.

So having showing that people fall more than is documented in claims is probably happening quite a bit. We probably are not seeing people who did not fall and having a claim for falls.

So what I think we're dealing with here is that a measure that's under-capturing the falls population and as we look to other data sources, we hopefully will be able to capture more of that population.

MEMBER WU: So, final question, do we think that there's variability in the ability of different organizations to capture falls? Because if it's really differential then we're going to have very unfair assessment.

DR. GIAVONETTI: I don't think there's differentiation at the health plan level. I think that the providers within the health plan are mixed

enough that there are some that are going to use the ICD codes for falls versus not using the ICD codes for falls.

Actually, what we can, and I don't have the numbers right in front of me, but I know we've looked at what the percentage of the population and the health plan fall into the denominator for this, the rate of falls, to see if we're seeing major variation in some plans identifying a lot more people with falls than other plans.

It's going to be influenced by the age of the people in the plans as well. And I can look at that and if you give me a little bit of time, I can go back and look at that. But that's also one indicator if we're seeing very different use of those codes across plans.

CO-CHAIR THRAEN: Any other questions for reliability? Shall we vote?

MS. QUINNONEZ: Voting is now open for the scientific acceptability of measure properties for reliability for Measure 2993. Option Number 1 is high, Option Number 2 is moderate, Option

Number 3 is low, and Option Number 4 is insufficient.

Okay. We're looking for two more votes. All votes are in and voting is now closed on reliability. The vote reads 45 percent voted high, 40 percent voted moderate, 15 percent voted low and zero for insufficient.

CO-CHAIR THRAEN: And just as a refresher, it's 65 percent I believe. Wasn't that -- 60 percent -- 60 percent.

MR. LYZENGA: So 60 percent, but which needs to be the combination of both high and moderate, those two. Yes. CO-CHAIR THRAEN: Just as a refresher, so.

MR. LYZENGA: If we do not reach the total of 60 percent in those two higher categories we're in the area of consensus not reached. That is if you don't get 60 percent in the lower categories. If you get somewhere in that gray area we're at consensus not reached status. And then we have a process for that as well, but I think we're in the clear on this one.

CO-CHAIR THRAEN: All right. Validity, Theresa?

MEMBER EDELSTEIN: Okay. Validity was tested at both the measure score and data element levels. Base validity and construct validity were both done.

They measured correlations with other measures of medication safety. They found moderate to high correlation between all rates except history of falls and chronic kidney disease with other medication safety measures.

The measure was deemed to have the attributes of a HEDIS measure. The geriatric measurement advisory panel at NCQA, its own committee on performance measurement, were both used in this process.

CO-CHAIR THRAEN: Albert? No. Questions? Shall we vote?

MS. QUINNONEZ: Voting is now open for the validity and scientific acceptability of measurement properties for Measure 2993. Option Number 1 is high, Option Number 2, moderate, Option

Number 3, low, and Option Number 4, insufficient.

We're looking for -- here we are. We have all votes. All votes are in and voting is now closed. The voting for validity of Measure 2993 reads 35 percent voted high, 45 percent voted moderate, 20 percent voted low, and 0 for insufficient.

CO-CHAIR THRAEN: Thank you.

MEMBER EDELSTEIN: Okay. Similar to the previous measure, administrative claims data are used for this measure. It's highly feasible.

CO-CHAIR THRAEN: Questions? Dr. Septimus?

CO-CHAIR SEPTIMUS: I've just been sitting here reflecting on this, this has got, what, four numerators? And we're not really sure about documenting falls and things, so I guess as I'm sort of sitting here is what's the feasibility of being able to capture all that with administrative data? It's a complex measure.

I'm not saying it's not important, but my question is I mean, based on your prior

experience of other plans, is this really easy and is it reliable and valid given all the shortcomings that other people on this committee have mentioned?

Now, maybe I'm misreading this, but that was my concern when I read this before I came up here.

MEMBER EDELSTEIN: So --

CO-CHAIR SEPTIMUS: Notice I didn't ask about age. But there is some differences in ages, by the way, based on the measure.

DR. GIAVONETTI: So, we think that in terms of feasibility, can a plan calculate this easily from their data, yes. They have all the data at their disposal around pharmacy claims and medical claims.

Going back to the reliability and validity of the falls indicator specifically, are we capturing everybody who had a fall? No. However, the population we are capturing we're very sure did have a fall and are receiving a medication. About half of them are receiving a medication that they should not be receiving.

So if the --- faced with the, either we don't measure it at all because we don't have a good way to get at falls, or we measure what we can which is those people where there is a claim for falls and half are getting a high-risk medication, we have chosen to measure what we can measure acknowledging that there needs to be improvement in the way falls are documented.

And we have other measures, that this panel has endorsed, that look at documentation of a history of falls, specifically two or more falls or fall with injury as recommended by the HES.

CO-CHAIR SEPTIMUS: Yeah, we had some misprint. So are there other measures around high-risk medications, and you'll have to education me on this, that get you the same information?

DR. GIAVONETTI: To our knowledge there is no measure that looks specifically at falls and the use of high-risk medication. There is a measure that looks at individuals with dementia in long term care facilities and their

receipt of antipsychotics.

Our measure looks at dementia across the entire plan population and looks at more than just antipsychotics. It also looks at benzodiazepines. So we are not aware of any measure that is getting at this particular package.

CO-CHAIR THRAEN: So, I'll get to you in a minute. So just to probably not stir the pot, or I probably shouldn't stir the pot, but I'm going to stir it anyway, is there any conversation about moving these cluster of measures towards a composite approach of any kind?

DR. GIAVONETTI: So that was a question that came to us from NQF about is this a composite measure or not. This measure was developed long before my time at NCQA, so I cannot speak to the people developing it. Did they think about composite measures?

It certainly was not tested as a composite measure. It was tested as, these are three populations where this is really important and we should look at the rate of -- and then,

wouldn't it be great if we just had a total that kind of gave you an overall sense.

So that's the way the measure is right now so that CMS or other programs can choose which rate they want to focus on. If they want one rate, they look at the total rate. If they want to look individually at the conditions, they look at those conditions.

It is a possibility that we could go back and look at our data and do some additional analysis to kind of understand more of the composite pieces and is there a different way to construct the composite, but we're not hearing particular feedback from any stakeholders that they'd like a different rate for this measure.

CO-CHAIR THRAEN: Lisa?

MEMBER MCGIFFERT: I may be sort of asking the same question that you were, but you said that there's a measure that looks at the accuracy of documenting falls, and -- no. No.

DR. GIAVONETTI: There is a measure that looks at whether or not individuals were

screened for a history of falls. And it says did you document in the medical record whether or not someone had a history of falls. It does not look specifically to say did you have a claims code then if they did have a fall.

CO-CHAIR THRAEN: So one of the things that occurs to me as we're talking about this, we're talking about fidelity as opposed to reliability and validity when you do an assessment, how well those assessments are being done. And that's a different kind of approach.

It's a fidelity question as opposed to a validity or reliability type of question I think, in the traditional sense of claims. You're looking at the process by which folks were being assessed and the data was being captured.

Other questions? Shall we vote?

MS. QUINNONEZ: Voting is now open for the feasibility of Measure 2993. Option Number 1 is high, Option Number 2 is moderate, Option Number 3 is low, and Option Number 4 is insufficient.

Okay. All votes are in and voting is

now closed. For the feasibility of Measure 2993, the vote reads 60 percent voted high, 25 percent for moderate, 15 percent low, and 0 for insufficient.

CO-CHAIR THRAEN: Next phase after feasibility is usability. Thank you.

MEMBER EDELSTEIN: Okay. This measure is already publically reported and used in accountability programs such as the health plan report cards, the health care annual report and the accreditation process for health plans as well as ACOs. So it's already in use in several ways for accountability and accreditation.

CO-CHAIR THRAEN: Ed?

CO-CHAIR SEPTIMUS: Quick question. If it's already being reported, and maybe I missed it, has it had an impact in terms of this measure? As the previous measure where we really have seen am impact, what about this measure that's being reported?

DR. GIAVONETTI: So I'm just looking. So this measure was updated in 2013, so we only have

two years of data. We did see a decrease in the dementia rate, looking at --- a small decrease. And I'm trying to find the chronic kidney disease. And also a small decrease in the chronic kidney disease rate.

MR. REHM: Yeah, I mean, I think in context, in our whole core of about 80 measures that we report through HEDIS, a half point or a three-quarters of a point is actually significant for measures that have played for a long time because, you know, the initial step is the first two years is great and then things get harder.

So I think the dementia rate actually decreased by a point and 1.2 percent at the mean. And so, that's a trend that we like. That's a good trend.

DR. GIAVONETTI: I'll also say this measure is not part of the financial stars calculation. So this one, it perhaps doesn't get as much attention by health plans, but we do hope that we are seeing trend in the right direction through its use in accreditation of their programs.

CO-CHAIR THRAEN: Any other questions? Pat, sorry.

MEMBER QUIGLEY: Thank you. Pat Quigley. My comment related to usability is still the focus on someone who has had one fall in the last year.

That fall could have been an accidental fall that happened because someone was walking and looking at the their phone, on their cell phone and had a distracted fall, the new type of fall, or it could've been those who have more than one fall, two falls, which then becomes a biological marker of maybe some underlying pathology.

I think the usability would have been better had this indicator coming forward as a new measure focused on those with more vulnerability as was the recommendation of the United States Preventive Services Task Force. Thank you.

CO-CHAIR THRAEN: Any other comments or questions? Vote, please.

MS. QUINNONEZ: Voting is now open for the usability and use of Measure 2993. Option

Number 1 is high, Option Number 2 moderate, Option Number 3 low, and Option Number 4 insufficient information.

All votes are in and voting is now closed. For the usability and use of Measure 2993 the vote reads 55 percent voted high, 35 percent voted moderate, 10 percent voted low, and 0 for insufficient information.

CO-CHAIR THRAEN: Next phase I think is endorsement overall. Uh-huh.

MS. QUINNONEZ: If there are no questions we'll move on to overall suitability for endorsement of Measure 2993. Option Number 1 is yes, Option Number 2 is no.

Okay. We're looking for one more vote. All votes are in. Voting is now closed. For the overall suitability and a recommendation for endorsement the vote reads 85 percent voted yes and 15 percent voted no.

CO-CHAIR THRAEN: All right. Hold on. Next one is 2988 Medication --

CO-CHAIR SEPTIMUS: Thank you, by the

way.

CO-CHAIR THRAEN: Thank you. And -- I'm sorry. Thank you.

UNIDENTIFIED SPEAKER: Thank you very much, folks.

CO-CHAIR THRAEN: I'm being rude. I'm being task oriented.

And we need to invite the Kidney Care Quality Alliance to the Table.

MR. LYZENGA: So, before we have our developers introduce this measure, I should mention we actually took this measure to our Renal Standing Committee, because it deals with patients who are in dialysis facilities. We figured we didn't think we had renal expertise on this committee. So -- and we had that committee seated already. So we just thought we would go and get a little bit of general input from them. They didn't vote on the measure or anything like that, just kind of gave us their thoughts and general feedback.

So just a few items that I pulled out

of that discussion and I'll allow our developers to sort of give their impression of that feedback as well. In general, the committee members were very supportive of the measure. They acknowledged that medication reconciliation is a very important issue for patients with ESRD. They -- those patients are often on multiple medications, ten or more, or -- have -- frequently have various comorbidities, diabetes, cardiovascular disease, and are seeing different providers who are giving them different medications and undergo frequent changes in their medication regimes. So reconciliation is a pretty important issue for this particular population.

Some sort of ideas that -- or things that they brought up about the measure themselves: some expressed some concern that it could be seen as a sort of check-box measure, given that the quality of reconciliation, itself, is not necessarily validated against the medications patients are actually taking.

The developer did note in that

conversation that this measure serves as more of a first step and that there are more comprehensive medication-review measures under development that take that more into account.

The committee members also talked about the fact that it can be find -- difficult to find a single source of proof when it comes to patients' medications. Those medications, again, because these patients are seeing different providers, those records can be disbursed across different sites and providers. The committee suggested that the patients, themselves, can often be the best source of information and suggested that future medication-reconciliation measures should incorporate some element of patient engagement, talking to the patients, themselves, or doing some sort of survey or something like that.

So that was -- that was generally the feedback from the Renal Committee and I'll turn it over to our developers to talk a little bit more about the measure and their thoughts on it.

DR. NISHIMI: Thanks, Andrew. I just

wanted, before Lisa summarizes the measure to you, to let you know that the Renal Committee, you know, gave us the feedback, obviously, on the measures. But I just wanted to let you know that this is actually one of three measures that Kidney Care Quality Alliance looked at.

We looked at a review measure -- sorry -- we looked at a review measure, a reconciliation measure and then a documentation measure. So this is -- this is kind of the middle step and it was felt by KCQA that it was the appropriate place to start and that the review measure, where there should and is more patient engagement, is definitely something that we're still looking at and developing. But to at least get something out on the table for this vulnerable population was important at this time.

CO-CHAIR THRAEN: Ed's asked that you all introduce yourselves to start, please.

DR. McGONIGAL: Okay. I'm Lisa McGonigal. I'm a consultant to the Kidney Care Quality Alliance.

DR. NISHIMI: And I'm Robyn Nishimi, also a consultant to KCQA.

DR. McGONIGAL: Okay. So, first, we wanted to thank you for taking the time to consider our measure here today. This is NQF2988, Medication Reconciliation for Patients Receiving Care at Dialysis Facilities, again, developed by the KCQA.

So the measure is specified at the level of the dialysis facility. It's applicable to all patients who are permanently assigned to a facility, this includes in-center patients, home patients, hemodialysis and peritoneal dialysis patients. It assesses the percentage of patient months for which medication reconciliation was performed and documented by an eligible professional.

In regards to importance, as Andrew just noted, medication management is a critical safety issue for all patients but is especially so for patients with end-stage renal disease. These individuals often require ten or more medications

and take an average of 17 to 25 doses per day. They usually have numerous comorbid conditions, multiple healthcare providers and prescribers, and they undergo frequent medication regimen changes.

Also, medication-related problems contribute significantly to the approximately 40 billion dollars in public and private funds that are spent annually on ESRD care in the United States.

So the measure is structured such that, rather than seeking a single yes or no check box that medication reconciliation was performed for a given patient in a given month, we require multiple elements must be met to be counted as a success on the measure. In addition to requiring that all known medications be reconciled by an eligible professional and the date of the reconciliation must be indicated, we also require that the identity of the eligible professional must be indicated and we specifically defined what must be addressed during the reconciliation process.

The measure was tested using data from

three KCQA member dialysis organizations, each with the capacity to provide retrospective analysis from a data warehouse or repository drawn directly from their electronic medical records. Testing involved approximately 5,292 facilities and this varied a little bit depending on the month. There were approximately 328,000 patients in each of the six months of the study, which was conducted from April through September of 2015.

The mean performance score during testing was 52.62 percent with a range of 0 to 100, meaning that some facilities did not perform medication reconciliation as defined by the measure for any patients. So there is significant room for improvement and a substantial gap on this.

Empiric testing was done using the beta-binomial method, which, again, is at the measure score level, and this demonstrated that the measure is highly reliable with a mean reliability of 0.9935 and that the measure components can be feasibly collected.

So I wanted to clarify two things

related to feasibility in our submission. First, as we noted in the submission documents, we identified a definition discrepancy among the three dialysis organizations as we developed the measure specifications. Specifically, while all three organizations that participated in testing identify and engage in the same three components of medication management, which is documentation, reconciliation and review, one organization flipped the definitions of reconciliation and review and put these in reverse to those detailed in the measure specifications. We'd note that this discrepancy was identified prior to measure testing, so that all three organizations used the same definition when they were testing the measure.

We also wanted to discuss unknown being an allowable response for some of the data elements required in the measure. So, as we indicated in the submission documents, depending on the electronic data system being used by a dialysis organization, some data elements can only be recorded in a free text field, which may or may not

have been completed. And, even when the data system does have a fixed field for a given element, the field may have been left blank. Thus, not all information sought in the measure is consistently readily available to the individual performing the reconciliation.

Our measure development work group noted, for example, that a dialysis facility personnel might have no way of knowing the precise start date or the particular clinical indication for a medication prescribed by another provider. So this issue necessitated that unknown be an allowable response for such irretrievable data elements so that, while facilities are expected to create the most accurate and complete reconciled list of a patient's medications possible, they are not unfairly penalized for not having access to information that they cannot reasonably be expected to have.

And here, again, I want to emphasize that this matter was identified prior to testing, so the testing organizations all approached the

specifications uniformly. And, with that, I would like to turn the floor over to the committee.

CO-CHAIR THRAEN: So the lead on this is -- hold on -- Missy. That's correct.

MEMBER DANFORTH: Thank you. So, first, thank you for the -- to the measure developers for doing a fantastic job describing the measure. So our first job as a committee, actually, is to review the evidence. I did have some concerns that I wanted to discuss. So first is a reminder this is a process measure. And, if you look at the measure framework for process measures, if a systematic review isn't included -- and in this case it was not -- I believe the measure can't be rated as a 1 for evidence. Can you confirm that, someone?

MR. LYZENGA: Yeah. That's -- yes.

MEMBER DANFORTH: Okay. So what the measure developer did do is actually provide a large body of evidence about the importance of medication reconciliation and reducing medication-related problems in ESRD patients.

However, I did actually review several of the studies that they included and almost all of them, actually, mentioned three components to really be effective in reducing medication-related problems in this particular group of patients and that was the reconciliation, the review, and then the management, not the reconciliation alone.

In addition, several of the articles and other articles related to this topic that weren't included do seem to suggest that the gold standard for med rec is -- that it's done by a pharmacist. And this measure actually includes a variety of eligible clinicians, including medical assistants and others.

And so I have -- I have, personally, just some concerns that the evidence that was provided doesn't really support the fact that this alone is going to have an impact on reducing medication-related problems in this group of patients, first, because it's a stand-alone reconciliation measure and the evidence really suggests that you need the reconciliation, the

review, and the management. And, second, because, again, the gold standard is that the med rec is performed by a pharmacist and not the other eligible clinicians that are included in the measure.

CO-CHAIR THRAEN: Response?

DR. McGONIGAL: Okay. First, I wanted to address -- and thank you for your comments. I wanted to address the gold standard. A gold standard, per se, has not been established and we wanted to point out that CMS specifically indicates for its Part D Medication Therapy Management Program that MTM services may be furnished by pharmacists or other qualified professionals.

CO-CHAIR THRAEN: Before you get off of that -- your list of qualified professionals so you -- CMS defines what those qualified professionals are. Is it the same list that you have that CMS defines?

DR. NISHIMI: I'm not sure about the CMS list. Our qualified professionals -- and it did vary depending on documentation,

reconciliation, and review, so, for this measure: physician, RN, an advanced practice RN, a PA, a pharmacist, or a pharmacy technician.

CO-CHAIR THRAEN: What's your understanding of the other issues?

DR. NISHIMI: The other issue for dialysis facilities is --

CO-CHAIR THRAEN: No pharmacists.

DR. NISHIMI: Yeah. The headquarters have --

CO-CHAIR THRAEN: There are none.

DR. NISHIMI: -- have pharmacy, you know, expertise at the sort of corporate level and we're aware that some are developing, you know, sort of teleconsultation services. But, in terms of doing a monthly reconciliation, there's just purely not enough people.

CO-CHAIR THRAEN: Did you have another -- or go ahead, Chris.

MEMBER COOK: As the lone pharmacist in the group, I would like to speak up on this. Missy, absolutely what you bring forth in the fact of med

reconciliation, it is the bare-bones minimum. And so, from a qualification standpoint, I think it's far more important that you're actually doing something on that. That's where even looking at a pharmacy technician or, you know, even an LPN or someone else who's actually looking because a med reconciliation is simply that you've got the list and you're making sure that it's the right drug, the right dose, and the right sig that's being associated with it.

You are not getting into the more cognitive points of ---

DR. NISHIMI: Management. Right.

MEMBER COOK: -- the management within that. Absolutely, that is what we need to go and I think, as a society and our system, we are headed in that direction but we are not there yet. And one of the great limiting factors is who is actually available to be there. And so, as a first start line of where we need to go to start raising that bar within that, I do support the fact of having med reconciliation because, as we see, so many of

the med error problems are in transitions of care.

And, when you have a highly vulnerable population, this is something that needs to be evaluated on a very regular basis because they see so many different doctors, such a large interval and they are taking so many medications. So it -- this is just the bare-bones minimum first start line for us to pass.

CO-CHAIR THRAEN: Lisa?

MEMBER MCGIFFERT: I just had a quick question about the description of the measure. So the measure is counting how many months in a year the patient got this reconciliation. So the expectation is that reconciliation happens every month. Is that correct?

DR. NISHIMI: Yes. So it's a patient-month construction and so, in month one, let's say a facility has a hundred patients. Did all 100 patients get a med rec? And so, if so, then they were 100 percent for that month. If, in month two they didn't -- none of them got it, then that would be zero percent. So, for a two-month period,

your score would be 50 percent.

It's a patient-month construction, not a patient because these patients are seen, generally, three times a week in center. And so, to do it on a percent-patient basis just didn't construct well.

CO-CHAIR THRAEN: And I actually, when I was reading this, made that thought to myself -- note to myself that the reason why there has to be a specific measure for kidney dialysis is actually the counting of the measurement itself, what you're going to count as opposed to length of stay or, you know, encounters. This is sort of a different animal and so it does need a little bit of a different methodology for counting.

Steve then Charlotte?

MEMBER LAWLESS: In terms of the importance of this, I was shocked by the gap. Fifty-two percent is just ungodly. But I also think that, to the point, this is a really raising of the bar. If you do medication reconciliation in a hospital, it is literally right now just

someone saying, do you take this, take this, take this? Ideally, it's all those elements that you've mentioned and they're -- most people are not doing it. And so this is really raising the bar, which creates some feasibility issues but I applaud this because of that.

The other aspect is the adverse-drug-reaction component of it. That's a bigger one. That takes a decision-making discussion about what that is and some evidence is about 22 percent of people have adverse drug reactions. So I -- that piece -- again, this is really raising the bar high. This may be raising the bar too, too high.

DR. NISHIMI: We looked at ADEs and thought about trying to construct the measure. But that was sort of beyond where we felt we could go because we were focused on even just documentation, reconciliation and review. So that's why we didn't march down the ADE path.

MEMBER LAWLESS: ADE's in there.

DR. NISHIMI: No. But a specific

measure only directed to ADEs, we decided not to do that.

MEMBER ALEXANDER: Well, I really appreciate the effort to address this population. It's one that I see and they're complicated and have so many doctors prescribing so many medicines with potential for problems. My concern is that this is not going deeply enough.

I do think that, when I see physicians and hospitals doing med rec, it is a mechanical matching. There's not the thought process that goes in: is there a duplication of medications that one doctor gave an anti-hypertensive, another doctor gave an anti-hypertensive?

So, unless you have the pharmacist or someone really critically looking to be sure that there's not a duplication, that there's not an inappropriate one, whether it's because of age or because of disease process, that it really misses where we want to go. And so I have a concern it's not going deeply enough.

CO-CHAIR SEPTIMUS: Getting back to

Steve's comment, I think we do need to raise the bar. Med rec is difficult enough, as we know, in the inpatient side and somewhat in the outpatient side, especially with the number of prescribers that are involved in the care of a dialysis patient.

The issue of lack of pharmacy support -- is that what you said there? If we were to pass this measure --

CO-CHAIR THRAEN: It would change.

CO-CHAIR SEPTIMUS: -- and we raise the bar and we provide safer care for dialysis patients, isn't that what we're trying to do? I hate --

I'm going to get a little bit on my soapbox and say that the regulatory lever over and over and over again drives care in a positive direction. There are sometimes some unintended consequences but, in a positive direction, because we don't do it ourselves. So, if it does push people to get the right people involved in these complicated patients, so be it.

CO-CHAIR THRAEN: The master has

spoken.

Yanling, go ahead.

MEMBER YU: Yeah, I definitely think that, just by one eligible profession to do the reconciliation, you know, sign up is a little problematic. I think, whether there should be a team to really have someone is lead pharmacist and then have, you know, cosigner, so that would make sure that, you know --

And, also, I have a -- I have a question about how do you -- how do you do it, this -- the feasibility of that. Is that later? How --

CO-CHAIR THRAEN: Okay. Pat?

And that's my phone. Ignore it.

MEMBER QUIGLEY: Thank you. And I appreciate all the comments, too, Pat Quigley -- related to the professional -- who is included in the professional. And, maybe, the consideration should be -- to the developer is that the pharmacy technician be excluded from the group who could do that or even an RN.

The person that would be doing the

medication reconciliation should be someone who has prescribing authority and understands the medications that someone is getting, rather than having the all-inclusive list, but to reduce the list of those who are able to complete this process so that it has a quality component, not just a check-box component.

CO-CHAIR THRAEN: So I don't disagree with anything that's been said. So I -- in my former life, I as in the regulatory end of the Department of Health and, you know, I'm from the government. I'm here to help.

And, currently, the -- and I want to verify this with you guys. Currently, the dialysis centers are wild, wild west. They operate pretty independently. They're not usually owned by the hospitals and we have challenges associated with regulating them because they are so independent. And, as a result, the criteria for the kind of staff you have to have is pretty minimal is my understanding. I mean we're talking about, basically, nursing staff.

And so that team that you're asking for I think is a great idea but it's not the reality, the current reality of how those systems are working today. And I also agree with the pharmacy but that's now the reality of how they're working today. So I think that you're coming up against the current regulatory environment in terms of what we allow for renal dialysis centers and how they operate.

Lisa and then -- and then Laura.

MEMBER MCGIFFERT: I think Yanling might have asked this but I'm also fairly familiar with the environment of a dialysis center and the limitations that is offered. My understanding is this is sitting down with the patient and going over what their understanding is of what they're taking, maybe bringing their meds in, in a bag or whatever you want to call it.

But I kind of like that about this measure because it is -- that's the patient's reality. And, yes, that patient might leave something off but, if they left it off, they're

leaving it off, probably. And that could be a significant issue. And so, you know, I -- what everybody said about then what happens is what we need to address next.

But it seems -- am I understanding correctly every month you'd sit down and go through with the patient what they're taking, correct?

DR. NISHIMI: Right. So you would -- that would be one component. You would also have, obviously, prescription that you may have prescribed. But, if the patient then brings -- doesn't bring them in, that raises a flag.

So to go to your point and the point that's made by others, the downstream sort of team environment review, that's the medication review measure that we did not bring to you and that has a smaller sphere of eligible professionals.

MEMBER ARDIZZONE: Thank you. Lisa, that was a great comment. I just wanted to respectfully disagree with Pat. I think it's well within a RN's scope of practice to understand what medications their patient is on and to reconcile

them. I can't speak to the pharmacy tech but I would assume they have some knowledge in pharmacology, so that they can review medication.

So I think those are important, eligible professionals, especially since we're talking about an environment where you -- it's not an acute care patient institution. You don't have six to seven providers for every patient, so you have to work with what you have.

MEMBER DANFORTH: Yes. Just to respond to the developer and everyone's comments. I definitely understand the importance of having medication safety measures for this group of patients and, certainly, that this setting can play an important role with that.

But, in sort of looking at the evidence that links this particular process to the stated outcome, which is reducing medication-related problems in this population of patients, the distance is further than other measures we have looked at. And so that's -- I think that's part of what we need to discuss is sort of how close is

the process to the outcome.

And, in this case, based on the evidence that was presented, it is to your point that combination of reconciliation, review, and management. And so, certainly, I understand the many challenges because I was here last year bringing composite measures to this committee. But, truly, I mean, this does seem like an opportunity to bring forward like a really strong composite measure that includes the three components that will actually then result in the strongest evidence that the measure will link to the outcome.

I'm just a little bit concerned that, you know, the lift of this measure is a little bit high because of the different -- it's a "and" and "and" and "and" in all these components. And so, to have, you know, people -- whatever the professionals are -- nurses at these centers doing this documentation, I'm not entirely convinced based on the evidence that you're going to see a reduction in medication-related problems is my point.

CO-CHAIR THRAEN: Charlotte, did you have another comment?

MEMBER GELINAS: Thank you. And I agree about the types of providers that can do medication reconciliation. I just -- when I read this measure, I want to applaud that we're even discussing it. When we get to the other -- the ambulatory SSI measure, I'll bet we have just as robust a conversation.

But the entire field of ambulatory care is, as you say, a bunch of cowboys. And we have to start somewhere. But we're talking accountability here and this will at least -- and Steve I agree with you a hundred percent -- you know, this is going to raise the bar.

But, at the end of the day, we are the Patient Safety Standing Committee. Our charge is to assure the public that we are doing our very best to improve patient safety. So I think it's exciting that we're even talking about this field of ambulatory. Let's not let the good be driven out by perfect and let's just hope that, in a few

years, that this measure will be sunsetted because it will have been so robust and we'll have reached a hundred percent, we've moved on to other measures.

But I do agree there are a number of providers that can do med rec. I would tell you that med rec, in general, is a wreck in healthcare today. The EHR providers, no matter where they are, inpatient or outpatient, certainly aren't helping this field. So, to whatever degree we can help clarify and amplify the importance, I think all of us around this table can offer expertise in that regard. But I do want to say bravo we're even discussing this whole realm of ambulatory. Thank you.

CO-CHAIR THRAEN: So we have Steve and then Kimberly.

MEMBER LAWLESS: Missy, actually to answer your point, if you look at the measure we just approved, the risk factor was kidney disease and falls in medications in one or two. That measure -- with medication reconciliation, that

measure would be cured because then we would say, you're on this measure and this measure, that caused a adverse drug reaction, which is a fall.

So, if you look at most of the measures we're reviewing today, anytime they mention kidney disease as a risk factor, most of this medication this is it. So this would be the process and those would be the outcome we would see. They would turn them into outcome measures and improvements. So just trying to connect the dots how crucial this could be.

MEMBER APPLEGATE: Yes. I just wanted to ask the group or the developers, if this measure overlaps at all with the last measure we voted on, 2993? There was some component about potentially harmful drug use in the elderly, at least with chronic renal disease. So we looked at potential opportunities for significant harm -- or harmful medication use.

And, when we looked at one of the components, the lowest use was less than ten percent of patients with chronic renal failure

received at least one harmful med. So that was the lowest group rate. And I just want to make sure that we're not doing any extra work or we're asking healthcare systems to do extra work with that metric -- overlap metric.

And, also, the other thing that I wanted to bring up was that we're asking healthcare systems to be accountable but we're also asking them to fix their electronic medical records and do things that I'm constantly asking them to do in the name of safety but without funding, so it's an unfunded mandate. You know? So just to address that. Thanks.

DR. NISHIMI: Do you want me to address that?

The measure is -- was feasible, highly feasible in the three dialysis organizations, large dialysis organizations that we tested in. It's -- they cover probably 70/75 percent of patients. So they have existing electronic clinical data streams in their facilities that feed into a corporate data warehouse. So they may have

to -- and some of the smaller ones may have to adjust some, no question.

But they don't have -- they collect this kind of data. As we said, some of it might be in free-text form that they may have to convert but, overall, we didn't have any complaints and they're not shy about expressing them in terms of that kind of burden now.

CO-CHAIR THRAEN: -- on. There. Some -- there are some substantial differences between this one and the last one. The last one was really aimed and very specific medication types and the relationship between that and age and disease state. And I think what your -- this measure's trying to get at is simply identifying what medications the patient's on, regardless of type. I don't think you're recommending that there by a judgment made on the type of medication that they're taking. I think there are some differences there.

Yanling?

MEMBER YU: Thank you. Just added to

what you said. I think that there -- personally, I do see the differences between this one and the last one. For one thing, the last one you -- basically, you looked up a pre-fixed table of medication -- what could be harmful medication; you identify them; and you score them.

But this one for dialysis patients, the reconciliation of medication also, hopefully, would help improve. You know, the medication could be duplicates, could be improper dosage and, you know, those type of things can cause harmful, you know, to the patient. And so I do see the difference.

Another thing I just -- I just wondering whether you would consider, because I really like with this you incorporate some patients' perspective or their knowledge into this whole process. I wish every medication reconciliation would do that but have you thought about to include documentation on whether risk and benefits has been clearly communicated with patients and the caregiver into this -- as a part of a medication

reconciliation?

DR. NISHIMI: We didn't. Everyone in the work group recognized that education was an important component. I think that we thought about it more in the context of the medication review measure, which is a little bit further downstream, more engagement with the patient trying to gauge whether the patient understood the fact that the reconciliation found differences. That's why it becomes a much more complicated measure and that's why, frankly, it's not before you right now.

This was a middle ground and, when we got the testing results and saw how poorly some facilities were doing, frankly, it seemed like we struck the right balance.

MEMBER YU: Okay. So, for your -- you have a list of allergies and all the, you know, adverse drug events experienced by the patient, does that mean patient reported or is that documented by a healthcare professional?

DR. NISHIMI: Both.

MEMBER YU: Both. Okay. So can be patient reported events. Okay. Thank you.

MEMBER COOK: Yes. I just want to point out -- and not to -- more for clarification because I think we've got to look at this and the expectations of it. And coming back to saying this truly is a first step. The reason that the list is broader of who can do this as a med reconciliation is literally just that you are looking at what drug and what drug. It's a listing and making sure it matches up.

It does not get to the cognitive component piece of managing the therapy or truly doing the review that gets into more of what you would have an advanced-practice nurse or a physician or a pharmacist truly go to do. So they're not probably going to -- may not catch those things that are Beers list. They're not going to catch those type of things.

It literally is patient comes in. They've got Zantac 75 and they've got Zantac 150. Why are you taking two drugs with two different

doses? Oh, we shouldn't do that. What should you be doing or -- you know, you're saying you're taking this once but, on the bottle, it says you should be taking this three times a day. It is literally at that form just of -- not from that therapeutic sense but just the operational sense of what's the list they should be taking and how they should be taking it and reconciling that to make sure it matches up.

CO-CHAIR THRAEN: So I see --

MEMBER COOK: It includes the adverse drug reactions, too.

MEMBER YU: Okay.

MEMBER COOK: Okay.

CO-CHAIR THRAEN: I see Kim's -- do you still -- did you have a question? No?

MEMBER APPLEGATE: No. I'm fine.

CO-CHAIR THRAEN: Ed? Come on, guys.

CO-CHAIR SEPTIMUS: Quick question. And, by the way, medical literacy is very important in all of this. I just want to follow up on Chris's comment. What's the age group that you were

talking about?

DR. NISHIMI: It was all patients.

CO-CHAIR SEPTIMUS: Okay. I just want -- I wanted to emphasize that. Previous measures were talking about us old folks. And, although, hemodialysis --

CO-CHAIR THRAEN: Speak for yourself.

CO-CHAIR SEPTIMUS: I am talking about myself.

But the reality is that, of course, some of these physiologically are much older than their stated age. But, nonetheless, we're talking about a much greater number of populations at a variety of different age and we're not excluding people because of their age. So this is a pretty broad measure, which I think is a good thing.

CO-CHAIR THRAEN: So I'm going to call for the vote.

MEMBER DANFORTH: Can I just ask a clarifying question?

CO-CHAIR THRAEN: Go ahead. Sure.

MEMBER DANFORTH: Actually, based on

people's comments I think I'm confused or maybe my initial understanding of the measure was not correct.

CO-CHAIR THRAEN: No. I think you did a critical review.

MEMBER DANFORTH: But my understanding of the measure is that you pass the measure if: you include the name of the eligible professional; the date of the reconciliation; address all known medications that are administered; for each of the medications, you have the name; and then you list any allergies, intolerance or list any adverse drug events.

There's no discussion of the adverse drug event. There's no discussion of why are you on two Zantacs; you should only be on one. Can you please clarify that? Maybe I misunderstood it.

DR. NISHIMI: No. That's -- the discussion measure is the review measure.

MEMBER DANFORTH: It's literally a list: I had an adverse drug reaction --

DR. NISHIMI: The medical record may

list it but, for purposes of the measure, it's, there was an adverse drug event.

CO-CHAIR THRAEN: So the challenge we have before us is the fact that I think there's consensus that this an important measure and an important first step. But the question is, does the evidence support that?

MR. LYZENGA: And, just to reiterate what Missy mentioned, because there wasn't a quality, quantity, and consistency explicitly stated in the -- of a systematic review, as stated in the submission, this is actually only eligible for moderate, at the highest rating. So start at 2.

MS. QUINNONEZ: Voting is now open for Measure 2988, Medication Reconciliation for Patients Receiving Care at Dialysis Facilities. We're now voting on evidence. Option Number 1 is -- Option Number 2 will be moderate, Option Number 3 will be low, and Option Number 4 will be insufficient. So your choices are: Option Number 2, moderate; Option Number 3, low; and Option

Number 4, insufficient.

We're looking for one more vote.

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

For the evidence of Measure 2988 -- we still had someone to vote for 1, for high. The vote reads 55 percent voted moderate, 35 percent voted low, and 5 percent voted insufficient.

Would you like to re-vote?

MR. LYZENGA: Can we recall the vote?

CO-CHAIR THRAEN: So a reminder you cannot vote for high in this situation. Your choices are moderate, low, and insufficient.

MR. LYZENGA: The next time we'll remove that from the voting slide, just to -- for clarification.

MS. QUINNONEZ: Okay. We're re-voting on Measure 2988 for evidence: Option Number 2, moderate; Option Number 3, low; and Option Number 4, insufficient. Click 2 for moderate, 3 for low, and 4 for insufficient.

All votes are in. Voting is now

closed. The vote reads 55 percent moderate, 40 percent low, 5 percent insufficient.

CO-CHAIR SEPTIMUS: That doesn't make any sense. If someone voted high, then you'd figure they would vote moderate.

(Off microphone comments.)

CO-CHAIR THRAEN: So we don't have consensus, so, therefore --

MR. LYZENGA: So we are at consensus not reached but we will move forward onto the remaining criteria, discuss all of those. We will revisit the measure after the comment period. We'll put the measure out for comment, see what kind of comments we get, revisit it in a vote and what we won't do is take an overall vote on this measure, an endorsement vote. We will vote on each of the subcriteria remaining and then we'll -- this will be consensus-not-reached status and then we'll see what happens during the comment period. We'll revisit it and then we have a process that goes forward from there.

CO-CHAIR SEPTIMUS: If you'll

remember, we had several measures that we've gone through this over the years. So --

MEMBER DANFORTH: So the developers, as they mentioned, did test the measure in three member organizations and there actually was a significant performance gap in the reconciliation, which they mentioned. I think the mean was 52 percent -- around 52 percent. I don't know if there's any other comments.

DR. NISHIMI: Yes, that's correct, and the range was 0 to 100.

CO-CHAIR THRAEN: I was a little bit confused about the standard deviation. Wide variability on that one. What are your thoughts?

DR. NISHIMI: Yes. Unfortunately, our methodologist isn't here. Is there a -- I'd have to --

CO-CHAIR THRAEN: When I was looking at the 95 percent confidence interval, for example -- and maybe this is my lack of statistical knowledge -- we're in the point something range for confidence but we have a standard deviation of

32.83. That didn't make sense to me. Am I not understanding?

MEMBER COOK: No, it doesn't make sense.

CO-CHAIR THRAEN: Thank you.

I went on to look at it and it was repeated in a couple of different areas. So, again, it just -- I was not making sense out of the -- out of the statistics.

DR. NISHIMI: Yes. We'd have to check because, as I said, the methodologist isn't here. But I see what you're saying. I think it's probably a typo.

MEMBER WU: So the comment is, if the standard deviation is really big, it's going to be very hard to detect changes, differences, or anything else using the measure.

CO-CHAIR THRAEN: -- other comments or questions about this?

Missy, did you have anything else you wanted to say about that?

MEMBER DANFORTH: The specifications,

themselves, were clear. The list of elements that had to be included were clear. There was a comment by one of the committee members about a lack of specificity around the definition for adverse drug reaction. I don't know if that individual wants to comment but, in general, the specifications, themselves, were clear and they did demonstrate a gap in performance.

CO-CHAIR THRAEN: Shall we vote?

MS. QUINNONEZ: Voting is now open for the importance to measure and report performance gaps for Measure 2988. Option Number 1 is high; Option Number 2, moderate; Option Number 3, low; and Option Number 4, insufficient.

Okay. We're looking for two more votes.

MR. LYZENGA: Michelle, could you vote for a performance gap? It looks like you voted for importance to measure.

MEMBER SCHREIBER: Okay. Sure.

MS. QUINNONEZ: We're looking for one more vote. Has everyone clicked?

All votes are in. Voting is now closed. For performance gaps of Measure 2988, the vote reads 35 percent high, 50 percent moderate, 5 percent low, and 10 percent insufficient.

CO-CHAIR THRAEN: Reliability.

MEMBER DANFORTH: So the developer included -- they did reliability testing at their performance score level and they actually had really significant results. The mean reliability score is a .99, which is extremely high.

They did not and I think someone else mentioned it that, if the measure got rolled out to a larger set of dialysis patients and dialysis centers than those that were tested, there would need to be some upgrade, probably, to the systems that the smaller dialysis centers were using that could impact the reliability.

All of these centers were members of this particular -- your collaborative. So I assume they have sort of a homogeneous level of --

DR. NISHIMI: No.

MEMBER DANFORTH: No? Can you speak

to that a little bit?

DR. NISHIMI: The -- there are three large dialysis organizations who own a range of types of facility and, as part of our membership KCQA and KCP's membership, we do have small independent dialysis facilities. It's just that, for testing purposes, because we could get a big population by using just these three -- and, frankly, others don't have as much bandwidth to help us out with testing because they're -- you know, they're just not as big corporately. So that's why we tested it there.

But we do have members, most of whom now have electronic systems. We just feel that, probably, their level of sophistication isn't as high, although they are electronic.

CO-CHAIR THRAEN: Yanling and then Albert.

MEMBER YU: Thank you. The question about the documentation identified for each medication -- there's a -- you can either document it or mark it as unknown, such as reason for

medication to -- stopped or discontinued. someone can put in "unknown." Is that a black mark or having anything to do with the overall score? I mean, if you stop medication, there have got to be some reasons. Could it -- it has to be documented. It seems like that.

DR. NISHIMI: If the prescriber was someone other than the dialysis facility, so if it was one of their physicians, the facility might not know when, exactly, it was stopped or why exactly it was stopped.

MEMBER YU: Then maybe somewhere should have a -- when items you have marked unknown, particularly like those type of situation it's important, and it should say, going to follow up or there -- something -- someone going to say something rather than just check it and say unknown and then that's it. That's sound like a little --

DR. NISHIMI: I think, frankly, the burden associated with following up every month for all patients, if you don't have an indication, would just be unreasonable.

CO-CHAIR THRAEN: Okay. So she's dumbfounded. We'll let you think about that a minute.

Albert?

MEMBER WU: -- a little more clarity on what test was done for testing reliability. The -- sort of the level of reliability that was cited was -- it was so high that it almost seems like, you know, simply if you turned your head and drank a cup of coffee, you know, you would make that level of -- create that level of unreliability. Can you just explain a little bit more? It seems almost too high to me.

DR. NISHIMI: It was a standard beta-binomial such as the one, you know, you just heard NCQA use, PCPI uses, signal-to-noise ratio. That's the way it came out.

CO-CHAIR THRAEN: So there's some concern about -- I think, because of the standard deviation question I had earlier and now you're kind of raising this other question statistically, I think this needs to be reconsidered or reviewed

to make sure that the numbers are correct.

CO-CHAIR SEPTIMUS: Can you -- let me ask you a question.

DR. NISHIMI: I know that reliability numbers are correct.

CO-CHAIR SEPTIMUS: Can you get that individual on the phone after lunch to respond or, otherwise, I can tell you this measure is in trouble?

DR. NISHIMI: I'll see what I can do.

CO-CHAIR SEPTIMUS: I mean would that be okay? I mean we can come back to it but I think, unless some of these other questions are answered, I think we're -- the Committee's having -- we're struggling here.

DR. NISHIMI: Yes. No. I know the reliability numbers are, because then we did a series of facility size, which is why we got to excluding less than 11. So those data I've looked at six ways and the reliability clearly goes down based on facility size.

CO-CHAIR SEPTIMUS: That's typical

overall then.

DR. NISHIMI: Right.

CO-CHAIR THRAEN: Kendall then Charlotte.

MEMBER WEBB: Did all three of the facilities you used have the same EMR?

DR. NISHIMI: No. And it's not three facilities, it's three dialysis organizations.

MEMBER WEBB: Okay. Do they use -- do they use similar EMRs. Because I know, like, in oncology, the same EMR is used across -- because it's just easier to collect data.

DR. NISHIMI: Not all -- well, the three large dialysis organizations have their own systems each. So then their facilities under the umbrella organizations would have the same.

MEMBER WEBB: Right. So did they have to put in something special in order to be able to say that they did these med recs?

DR. NISHIMI: No. They -- they are already collecting these datas.

MEMBER WEBB: Okay.

DR. NISHIMI: They have their own internal.

MEMBER WEBB: But they -- it's their own internal EMR, it's not the big five?

DR. NISHIMI: Correct.

MEMBER ALEXANDER: Did you look at the number of unknowns that were recorded and -- to see what size or percentage it was of all the data that was reported?

DR. NISHIMI: No.

CO-CHAIR THRAEN: Any other questions?

MEMBER WU: While we're on that question -- and, perhaps, questioning a, you know, sort of consultant-backed -- wherever that person is sitting -- I'm curious of whether or not unknown -- an unknown was classified as being agreement or how that was handled?

DR. NISHIMI: I'm not sure.

MEMBER WU: If two pieces of information were classified as unknown, were they -- were they thought to be in agreement with each other? Was that scored? How was that handled?

DR. NISHIMI: So, if one list had "unknown" for the indication and the other had "unknown," are they considered --

MEMBER WU: Is that considered perfect agreement?

DR. NISHIMI: All they're doing is, yes, reconciling the two. So, yes, it's a match. The source of the unknown would be handled under the review measure.

CO-CHAIR THRAEN: It's back to how they judge, that, right?

DR. NISHIMI: Yes.

CO-CHAIR THRAEN: As opposed to, in the data field, you have data field 1 has "unknown" and data field 2 has "unknown" --

DR. NISHIMI: Correct.

CO-CHAIR THRAEN: -- there's no way of -- other than what's in the data field.

How much of what you -- what was analyzed was text base?

DR. NISHIMI: We didn't analyze the actual text base.

CO-CHAIR THRAEN: Other questions or concerns about this? Shall we vote on reliability?

MS. QUINNONEZ: Voting is now open for reliability of Measure 2988: 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 reliability of Measure 2988, 15 percent voted high, 35 percent voted moderate, 30 percent voted low and 20 percent insufficient.

CO-CHAIR THRAEN: This does not pass.

MR. LYZENGA: Yes. So we're in the consensus-not-reached area again, so we will continue moving forward again. This will be another one that we'll revisit after the comment period.

MEMBER WU: Yes. I would comment that my vote was -- should really have been "unknown." And that's not --

CO-CHAIR THRAEN: Albert.

MEMBER WU: No. But that's not a --

that's not a strike against the proposal. I, you know, would be able to make a better judgment and, perhaps, a more favorable judgment if I had a bit more information.

CO-CHAIR THRAEN: Validity, Missy?

MEMBER DANFORTH: Okay. So, for this measure, the measure developer did phase validity only. They basically brought together two different sets of experts. One was a set of experts from the ER -- the end-stage-renal-disease field. The other was a set of experts -- I'm sorry. I'm just trying to scan this quickly.

CO-CHAIR THRAEN: There were two groups?

MEMBER DANFORTH: Yes, there were two groups and they basically asked two questions: 1) how likely is the measure score -- how likely is it that the measure score provides an accurate reflection of medication reconciliation; 2) what is the likelihood that the measure can be used to distinguish good from poor quality? Both groups said likely or highly likely at least 77 percent

of the time.

So the KCQA member organization group rated the measure 77.3 percent, so the measure would be likely or highly likely to provide an accurate reflect of med rec. And then the same 77.3 percent of the panel agreed that the measure would be likely or highly likely that the measure can be used to distinguish good from poor quality.

The expert panel had just slightly higher agreement, so 88.9 percent of the nine-panel expert panel said that the measure would result -- would be likely or highly likely that the measure scores reflected accurate med rec. And then 77.8 percent of that expert panel agreed that the measure would be highly likely or likely to be able to distinguish good from poor quality.

CO-CHAIR THRAEN: Questions? Comments?

Go ahead Yanling.

MEMBER YU: There's a question about this -- the performance gap, you know, the uncertainty. So I don't know how to vote on this

one because they have to go check the numbers to make sure. You know what I'm -- what I mean?

CO-CHAIR THRAEN: So you're raising the question about the standard deviation here?

MEMBER YU: Yes. That's --

CO-CHAIR THRAEN: Okay. So Yanling's saying, because there's a question on the standard deviation question, she's uncomfortable voting on the validity of the -- of the measure. Any other thoughts or concerns about that issue? You could invoke --

DR. PINES: I just had --

CO-CHAIR THRAEN: Go ahead.

DR. PINES: Just a comment. I think the validity data was provided and the issue was the standard deviation of the reliability, which we've already voted on.

MEMBER DANFORTH: Yes. And I would think, because they provided information on phase validity and not construct validity, that we could still look at the phase validity information they provided.

MR. LYZENGA: Although, I should note again that, according to our algorithm here, phase-validity-only testing means that the ceiling for this measure is moderate for validity ratings. So moderate, low, and insufficient will be our options.

CO-CHAIR THRAEN: So 2, 3, and 4 are your only options to vote on this one. So we call a vote.

MS. QUINNONEZ: We are now voting on the validity of Measure 2988. Your options are: Option Number 2, moderate; Option Number 3, low; and Option Number 4, insufficient -- Option Number 2, moderate; Option Number 3, low; and Option Number 4, insufficient.

All votes are in and voting is now closed. The vote reads 55 percent voted moderate, 35 percent voted low, 10 percent voted insufficient.

MR. LYZENGA: All right. Yet another in the gray zone here that, once again, we'll revisit. But move on to the next criteria.

CO-CHAIR THRAEN: Usability. Feasibility -- I missed one. Yes.

MEMBER DANFORTH: So, as the measure developers discussed throughout the discussion, the measure was tested in three large centers. If you look at the size of the sample, it actually includes a large sample of patients across the three testing facilities. They also address some of the known feasibility issues around definitional confusions and dealing with unknowns and clarify that they have adjusted for both.

CO-CHAIR THRAEN: Questions or comments?

Take vote.

MS. QUINNONEZ: We are now voting on the feasibility of Measure 2988: Option Number 1, high; Option Number 2, moderate; Option Number 3, low; and Option Number 4, insufficient.

Okay. We're looking for one more vote. Can everyone resubmit their clicks one time for me, please, pointing this way?

All votes are in. Voting is now

closed. The vote for feasibility of Measure 2988 reads: 30 percent voted high, 55 percent voted moderate; 5 percent voted low; and 10 percent insufficient.

CO-CHAIR THRAEN: Usability.

MEMBER DANFORTH: So the measure is not currently used in any public reporting or accountability programs but the developer did note that it is -- they'd like to see it used in the future in an accountability program and, also, that variations of the measure are currently in use by a number of dialysis organizations for internal quality improvement.

CO-CHAIR THRAEN: Yanling, go ahead.

MEMBER YU: You mentioned that there's a plan where you include public reporting and a payment program. And I wonder if the developer has any ideas or any thoughts, could share how you incorporate this type of a measure into the --

DR. NISHIMI: Yes. In fact, since we submitted the measure submission, the dialysis facilities are paid under a PPS system and then

what's called the Quality Incentive Program, QIP. And so, in the proposed rule that was issued on June 30th, CMS has -- because they knew we were developing this, has asked for comment on what the broader community, obviously beyond this, thinks about medication -- incorporating a medication reconciliation in the QIP, which is a penalty-based performance system.

So going forward, we do anticipate they would pick it up because they were part of, you know, following the development.

MEMBER YU: So there would be any comparisons on the facility level to this --

DR. NISHIMI: So it would be part of -- it would be facility-to-facility public reporting and, then the way the QIP is structured, their payment is based on that. They can be penalized up to two percent across the total performance score.

MEMBER YU: Would there be an example for other type of medication reconciliation measures to come -- to follow?

DR. NISHIMI: You'd have to ask CMS that.

MEMBER YU: Okay.

CO-CHAIR THRAEN: Probably.

MEMBER YU: All right. Thank you.

CO-CHAIR THRAEN: Probably.

Missy, go ahead.

MEMBER DANFORTH: We're going to talk about the competing measures at some point, right, but not now?

MR. LYZENGA: Yes.

MEMBER DANFORTH: Just to Yanling's point?

CO-CHAIR THRAEN: Go ahead.

MR. LYZENGA: We can talk about that right after this discussion on this measure.

CO-CHAIR THRAEN: Other questions or comments before we go into competing measures?

Go ahead, Missy.

Do you want to -- do we vote first on usability?

MR. LYZENGA: Yes.

CO-CHAIR THRAEN: Okay. Vote first on usability.

MS. QUINNONEZ: Voting is now open for the usability and use of Measure 2988. Option Number 1 is high; Option Number 2, moderate; Option Number 3, low; and Option Number 4, insufficient information.

And we're looking for two more votes.

All votes are in and voting is now closed. The vote for usability and use of 2988 is: 25 percent voted high; 60 percent voted moderate; 15 percent, low; and 0 percent for insufficient information.

MEMBER DANFORTH: I will try to do competing measures. I've never done competing measures before.

MR. LYZENGA: I think -- I think we'll actually want to take a vote on this, first, on overall vote on this measure.

CO-CHAIR THRAEN: Well, we don't endorse this one.

MEMBER DANFORTH: No.

MR. LYZENGA: Oh, right. We're not -- we're not doing overall. My fault. Yes, so now we can go into the related and competing discussion. So --

MS. MUNTHALI: But, just to clarify, we wouldn't include this one in the related and competing --

MR. LYZENGA: Because --

MS. MUNTHALI: -- because you haven't rendered a vote on it.

CO-CHAIR SEPTIMUS: That's true.

MS. MUNTHALI: But on the other two, you would.

CO-CHAIR SEPTIMUS: That's true.

MS. MUNTHALI: So, perhaps, you'd want to do that on a post-comment call or something?

MR. LYZENGA: I think that's a good idea. It makes sense.

MEMBER MCGIFFERT: Can someone explain this a little bit more?

MR. LYZENGA: So --

CO-CHAIR THRAEN: So this is a

non-endorsed one because it didn't meet the earlier criteria.

MR. LYZENGA: It was -- the measure is not officially endorsed. Usually, we do the related and competing discussion after we do that overall vote, because if a measure does not get a recommendation for endorsement, it's kind of a, you know, moot point. It's not related and competing with any endorsed measures because it's not endorsed itself.

Here, we didn't do that overall vote. So we don't -- have not rendered a final decision. So maybe we should wait on that question of related and competing until we have, in fact, rendered a final decision and see that this measure is --

MEMBER DANFORTH: Okay.

MR. LYZENGA: -- recommended for endorsement. Does that make sense to everybody?

CO-CHAIR THRAEN: Martha?

MEMBER DEED: Yes. I just had a comment that, when a measure looks as important as this one could be, I just think -- and it's kind

of a beginning effort, I think it's really, really important to have the present -- have the material utterly, utterly clean, so that we can work on it because this is a measure that I think might set a standard for some other work to come. So I would just really encourage the developers to, you know, just check things out and, hopefully if there's time now, today or tomorrow, to have -- revisit it, at least, you know, as a kind of a starting point to continue this discussion, because this is an important, important measure.

CO-CHAIR THRAEN: Yes. To follow up to that, if we're unable to do the clarification on the statistics while we're still here, we can do that in a follow-up call, making sure that all of the data is correct and our statistical questions are answered. And then we can actually do a -- an endorsement at that point as well. So we have a couple options.

I think everybody's in agreement. We've been in this place before where everybody's in agreement that this is a vital component but it

just hasn't met the scientific criterion. And so we have to -- we have to be true to that process and ask the developers to do a little more homework for us.

DR. NISHIMI: Right. And I have emailed them to check with the standard deviation. I am quite confident about the reliability statistics but I don't know about the standard deviation.

CO-CHAIR THRAEN: So I think we're done with --

DR. McGONIGAL: Thank you.

CO-CHAIR THRAEN: Okay. So we --

Public comment, anybody on the phone or anybody here that cared to make a public statement?

OPERATOR: Okay. At this time, if you'd like to make a public comment, please press star and then the number one.

Okay. And, at this time, there are no public comments from the phone line.

CO-CHAIR THRAEN: Okay. So, given the fact that we had 15 minutes for public comment, I

just want to note that we're on time. So I think we're taking a lunch break, aren't we? Yes. Good job, guys.

CO-CHAIR SEPTIMUS: All right. So we're supposed to be back at 1:15.

(Whereupon, the above-entitled matter went off the record at 12:03 p.m. and resumed at 12:46 p.m.)

CO-CHAIR SEPTIMUS: Okay. I want to thank our developers for responding quickly so we could get some of those statistical questions answered, so that we can make sure we get a full and honest evaluation for the measure that we took up at the end of the morning. So I'll let them, perhaps, introduce our folks -- your folks on the phone.

DR. NISHIMI: Is the phone line open?

OPERATOR: Yes, the phone line is open.

DR. SCHNEIDER: Thank you.

DR. NISHIMI: Craig, are you on?

DR. SCHNEIDER: Yes. I just -- I just dialed in.

DR. NISHIMI: So, Craig, this is the NQF Patient Safety Committee and they have some questions about the standard deviation for the med rec measure and the standard error. Did you want him to just --

CO-CHAIR SEPTIMUS: You can discuss it and then, anybody else who has questions about some of the statistics and validation, this is the time to answer those questions. So we'll take -- we'll take now until, you know, perhaps at 1:00 and then we'll decide whether or not we want to reconsider the measure, if that's okay. Okay? So go for it.

DR. McGONIGAL: Craig?

DR. SCHNEIDER: Yes.

DR. NISHIMI: Go ahead.

DR. SCHNEIDER: Okay. So there were some questions about why the standard deviation was so large and the standard error was so small. Is that correct?

DR. NISHIMI: Yes.

DR. SCHNEIDER: Okay. So the standard deviation in this case -- so the distribution of

the individual observations ran from 0 to 100 and it was kind of a U-shaped distribution. So there was a number at 0 and low and there was a number at 100 and high. So there is a large spread and that results in a large standard deviation, which, again, is a measure of the spread of the individual observations.

To get the standard error, what you do is you take the standard deviation and you divide it by the square root of the sample size. So, if you have a large sample size, then the denominator is going to result in a small standard error. But, even if you have a large standard deviation, if you have a large sample, then you will get a small standard error.

And that standard error -- keep in mind that the standard deviation and the standard error are also measuring two different things. So the standard deviation is about the individual observations where the standard error is a measure of uncertainty of a single number and that number is the mean.

And so, really, the standard error is used to create a confidence interval of the mean, which is not a confidence interval of where everyone will lie but it's a confidence interval of, if we were to do this again and get everyone's measure and an average, where do we think that single number is likely to lie? So that's why you see a large difference between a standard deviation and standard error.

CO-CHAIR THRAEN: Great. Thank you. And then we had some questions about reliability.

MEMBER WU: Could I --

DR. SCHNEIDER: Sure.

MEMBER WU: Could I just ask a question about --

CO-CHAIR THRAEN: Go ahead.

MEMBER WU: -- if you said the distribution is very not normal, is it -- is it appropriate to be reporting a mean score at all?

DR. SCHNEIDER: It's fine to report a mean. And that's -- the mean is still the mean regardless of if it's U-shaped. If you have a

strongly skewed distribution, whereas you have a lot on one end and not a lot on the other end, some would argued that the median is more appropriate than the mean.

But, if it's relatively symmetric, the mean and the median are going to be relatively similar. And, again, all of these measures are just -- they're summary measures but none of them individually tell the whole story right.

CO-CHAIR THRAEN: Other questions? Can we scroll down to the reliability section? I'm not generating the question we had in my head from before. So I need a queuing.

DR. NISHIMI: The question was why was the reliability number so high.

CO-CHAIR THRAEN: Oh, yes, on the binomial -- on the binomial --

MR. LYZENGA: And maybe just some clarification on the beta-binomial method and --

CO-CHAIR THRAEN: Yes. Correct. Thank you.

MR. LYZENGA: -- what it suggests.

DR. SCHNEIDER: Clarification on what the beta-bimanual is or why it was used or --

MR. LYZENGA: I don't know. Al, is that fair to ask? Is that what you were looking for, what the -- what the methodology was and why it was --

MEMBER WU: I was also curious to see that the coefficient that came out was almost perfect and that seemed very good but, potentially, you know, could have -- could have been an error, also.

DR. SCHNEIDER: Well, we -- so, in terms of -- both myself and Dr. Gilbertson ran the reliability independently and produced the same numbers. So we're confident in its accuracy.

In terms of the numbers themselves, I'm not sure -- I'm sorry. I'm not -- I'm not sure I'm totally understanding. So we're uncertain about or we had some confusion about the number itself or its value?

CO-CHAIR THRAEN: I think the original question was the size of the value.

So could we go down to the numbers? There they are.

So we have .9935 as the mean reliability of the measure.

DR. SCHNEIDER: Yes.

CO-CHAIR THRAEN: And that was perceived to be so close to one that it seemed too good to be true.

DR. SCHNEIDER: Oh, okay. So the -- in the beta-binomial -- in the beta-binomial, you actually get a reliability value for each individual facility because, in the beta-binomial, the actual performance -- how well you do is part of -- well, in the calculation of the reliability. So reliability is at least in part a function of the performance. And that is just simply because of the underlying distribution that's assumed when this is -- this is performed.

So the fact that the mean is as high as it is means that there were a lot of people with the reliability of 1 or very close to it. So that's, again, simply a function of the performance

of individual facilities and how they did and how many patients were included at each facility. And so that's the number that we obtained and I don't know how else to sort of -- sort of say it. It is high but it's -- in other ones that I've looked at, it's not unheard of and, like I said, it was run independently by two different -- by myself and Dr. Gilbertson. So we're confident in its -- in its accuracy.

CO-CHAIR THRAEN: So I just -- for question purposes -- so you said that we had a U-shaped curve, meaning there were a number of institutions that were close to zero.

DR. SCHNEIDER: Yes.

CO-CHAIR THRAEN: And a number of institutions that were close to 100 percent. So, when you did this reliability testing, is it on those facilities that are non-zero facilities?

DR. SCHNEIDER: That's on everybody and I will -- and I will say that, if you have a zero -- it's really -- the closer you are to either extreme --

CO-CHAIR THRAEN: The more reliable?

DR. SCHNEIDER: -- the more -- the more reliable you are. So having a huge distribution, in some ways, enhances that reliability --

CO-CHAIR THRAEN: Got it.

DR. SCHNEIDER: -- because there's less -- I mean, if you have a zero out of -- that's just inherently more reliable according to the method than, say, 50 percent.

CO-CHAIR THRAEN: Okay.

MR. LYZENGA: I thought, maybe, we should also note that the measures we just passed from NCQA also used the same methodology, the beta-binomial, for reliability and had very similar scores, 97 or above -- thereabouts.

CO-CHAIR THRAEN: So, given these explanations, are there any other questions about the statistics? Do we want to re-vote?

MEMBER WU: So I'll ask another question, which actually isn't for the statistician so much as for -- maybe for another explanation of what was actually being done in the

facilities. I know that, when I try to get anyone to reconcile anything in the clinic, people don't agree. So I am -- you know, so I wanted to -- some clarification for what was -- what the task was that was actually being done that yielded such high agreement among raters, perhaps, if that's what was happening?

DR. NISHIMI: It's not independent raters. It's someone in the facility performing the reconciliation function. Did they do that? Did -- was it documented? Did -- you know, did they check through the various elements: indication, the med, dosage, frequency, et cetera? Did they do that that month? Is it documented that -- the date of which they did it and is it personally identified to that individual, so not just a yes, someone did it but they either used their name or their employee-identifier?

If all three are present, then, for that patient there is a success. If all three elements are not present, then for that patient in that month it's a failure. So, for the facilities that had

scores of 100, that meant that, in that given month, they did all the elements of the reconciliation for every patient.

In the facilities that had a zero and -- it meant, for that month, they didn't do a reconciliation for any patient in that month. And so then it was a six-month test period. So then those who still had a zero meant that they did not do a reconciliation for any patient for any month.

CO-CHAIR THRAEN: Steve?

MEMBER LAWLESS: That piece, you're saying there's four elements --

DR. NISHIMI: Three.

MEMBER LAWLESS: -- or three elements, it's not. There's a lot of -- if there was 11 medicines --

DR. NISHIMI: Well, yes.

MEMBER LAWLESS: -- and 17 different measurements thereof to the reliability of do I have to get all elements of the medication correct on every medicine in order to be --

CO-CHAIR THRAEN: A hundred?

MEMBER LAWLESS: -- a hundred percent?

DR. NISHIMI: To meet that part of it, yes.

MEMBER LAWLESS: Yes. So the three -- third category or whatever --

DR. NISHIMI: The third category.

MEMBER LAWLESS: -- has to have --

DR. NISHIMI: They have to match.

MEMBER LAWLESS: -- 77 things correct?

DR. NISHIMI: Right.

MEMBER LAWLESS: And there's reliability on that piece to someone else saying the same thing?

DR. NISHIMI: No, it's not -- it's not inter-rater reliability.

MEMBER LAWLESS: Okay.

DR. NISHIMI: It's the reliability that the process was performed. You're not comparing two people at a facility doing it.

MR. LYZENGA: Yes. I should note this is not data element reliability.

DR. NISHIMI: Right.

MR. LYZENGA: It's reliability of the measure score and, maybe it's foolish for me to start talking about this because I'm not a statistician and I'm not that familiar with it but, as I understand it, it's a way of looking at, again, the signal-to-noise at the performance-score level.

And what you're looking at is the variability within an institution, within facility, that's being measured and then looking against the variability across institutions and trying -- and using a sort of statistical method of seeing how much of the variation in performance is due to that sort of variability within an institution, which is kind of what you call noise, and then true variation across institutions, which is, you know, the signal so to speak.

It's not really looking at whether they have done, you know, these things; what they're supposed to do as part of the measure but, in some sense, really looking at the ability of the measure to distinguish between facilities' performance.

And, again, I probably --

DR. SCHNEIDER: This is Craig. I -- this is Craig. I couldn't have said it better myself. That's exactly what this is doing. It's not inter-rater, it's not test/retest reliability. This -- but what you described is exactly what this is.

CO-CHAIR THRAEN: Laura?

MEMBER ARDIZZONE: I had thought maybe you had asked a question if we were -- wanted to vote again and I wanted to say, yes.

CO-CHAIR THRAEN: Well, before we vote, any other -- any other comments or questions?

CO-CHAIR SEPTIMUS: Just so we're clear, the signal-to-noise ratio was?

DR. NISHIMI: .99.

CO-CHAIR SEPTIMUS: .99, which we all say is extremely good.

MR. LYZENGA: That means virtually all of the performance -- variation in performance is due to real variation between facilities.

CO-CHAIR SEPTIMUS: Correct, which is

what we'd like to see. And, once again, give us the three elements that they have to have to meet the measure.

DR. NISHIMI: Perform the reconciliation. That includes --

CO-CHAIR SEPTIMUS: Right.

DR. NISHIMI: -- those components, date, and an identifiable individual who's an eligible professional.

CO-CHAIR SEPTIMUS: Okay. I just wanted to make sure everybody gets those three. Okay.

DR. PINES: And we also wanted to clarify exactly which votes that the Committee wanted to re-vote on, just to -- just to make it clear. So the -- so the evidence was consensus-not-reached. The performance passed -- performance gapped passed. Reliability and validity also were consensus-not-reached.

CO-CHAIR THRAEN: So, in the original conversation at -- the first one in terms of consensus-not-reached, I think the statistics that

we had questions about informed that decision. So I'm thinking that we need to go back and rethink that decision.

It also had to do with the -- Missy's critical analysis of the fact that the literature and the research that was provided in the documentation looked at those three components, which was reconciliation, review, and management and that this particular measure -- that literature did not support the single-case use of reconciliation. I think that the was the concern that you had is the disconnect in the science to support that one component.

That being said, the association is working on the other two-measure components in a -- for future work. So there's an anticipation that those measures might be brought back once the work is done on testing and validating those two measures.

MEMBER WU: Could -- could I ask the two of you or one of the two of you to convince me and us that this -- doing this component of what is

being done will -- is, itself, a useful -- a valid exercise and a useful measure that we really ought to be grading and maybe even rewarding or punishing people about?

DR. NISHIMI: I think that this patient population is one of the sickest populations in -- receiving healthcare right now. They have a lot of meds. They're seen -- if they're in-center, they're seen three times a week, sometimes four, depending on the patient, in the facility. So their meds -- and then they're also seen by a physician usually at least -- at least once a month, often more often and then they may be seen in an outpatient capacity by a cardiologist, not the nephrologist.

So they have a lot of touches but the main touch is at the dialysis facility. So that's the real opportunity to get them together and see -- I don't know if that's a good comment or a bad comment on the line.

(Off microphone comments.)

DR. NISHIMI: Oh, okay.

They have the opportunity in the facility to perform what we believe is a critical function, which is, granted, only the beginning of a multi-stage process. But it -- you have to start somewhere and the fact that you have, you know, so many facilities in a U-shaped curve, frankly, not even doing a single med rec for a six-month period I think is just astounding. That's a polite way to put it.

So that's why we think it is -- I mean we -- there is no systematic review and there is no single study looking just at the med rec component. But medication management is clearly important for this population and so you have to start somewhere when you start measuring this stuff.

CO-CHAIR THRAEN: Chris?

MEMBER COOK: Again, I'll come back as the pharmacist within that. There is a tremendous amount of literature that looks at the general population. When you're talking about from med reconciliation, from review, when you get into med

management, there's a lot of stuff that's building. The hard part about a lot of it is this is a process measure. This is a hope to avert situation.

And, as a practitioner, the scary part to me on this is this is something you should be doing that's very basic. And the really scary part is it's not happening. And so it's almost absurd that we have to ask for these things. It's almost like, wow, did you wash your hands before surgery? But we saw that it was -- we have some issues there.

This is very much, in the medication standpoint, that very same similar piece. You have very complex patients. You have a lot of drugs with a lot of different physicians. You need to at least have the very basic stuff of reconciliation to begin that and then, hopefully, you're able to start pushing more to where we can get to that more advanced stuff, review and management, down the line. But it is a critical first step.

CO-CHAIR THRAEN: Any other -- who's that there? Oh, no. Laura, she called for the

vote. That's what hers was off of.

Any other conversations or comments, questions? So we're going to call for the vote to start at the beginning, which is the --

CO-CHAIR SEPTIMUS: Well, the ones we didn't get to consensus on.

CO-CHAIR THRAEN: Yes. Yes. You have that up already? Okay. Thank you.

MS. QUINNONEZ: Yes. We are now re-voting on the evidence for Measure 2988. You have three options: Option Number 2 is moderate; Option Number 3 is low; and Option Number 4 is insufficient. Option Number 2 is moderate; Options Number 3, low; and Option Number 4, insufficient.

Got it. All right. We have all of our votes. Voting is now closed. The vote for the evidence of Measure 2988 is 74 percent moderate, 21 percent low, 5 percent insufficient.

CO-CHAIR THRAEN: It's passed. Now, we're going to go to reliability, correct, or performance? No. We passed performance.

Reliability.

MS. QUINNONEZ: Okay. We're voting for the reliability of Measure 2988: Option Number 1, high; Option Number 2, moderate; Option Number 3, low; and Option Number 4, insufficient. You may cast your votes. Option Number 1, high; Option Number 2, moderate; Option Number 3, low; Option Number 4, insufficient.

We're looking for one more vote.

All votes are in. Voting is now closed. The vote for reliability of Measure 2988 is 47 percent high, 53 moderate, 0 percent low and zero percent insufficient.

MR. LYZENGA: I think we no longer have consensus not reached on any of the --

CO-CHAIR SEPTIMUS: Yes. So now we can go to whether or not we want to --

MR. LYZENGA: Oh, validity, too. Sorry.

CO-CHAIR SEPTIMUS: Validity, too. Okay.

MS. QUINNONEZ: Voting is now open for

the validity of Measure 2988: Option Number 1, high --

MR. LYZENGA: Sorry.

MS. QUINNONEZ: Sorry. Three options.

MR. LYZENGA: So, again, for this one, there's only phase validity, so eligible for moderate at the --

MS. QUINNONEZ: Got it. Here we are. For the re-vote of Measure 2988, we have three options: Option Number 2, moderate; Option Number 3, low; and Option Number 4, insufficient. For the validity of Measure 2988, Option Number 2, moderate; Number 3, low; and 4, insufficient.

All votes are in. Voting is now closed. The vote for the validity of Measure 2988: 89 percent moderate; 11 percent low; 0 percent insufficient.

CO-CHAIR SEPTIMUS: Now we go is the measure suitable for endorsement, which we did not vote on because we didn't reach consensus. So this is the last question for this measure.

MS. QUINNONEZ: Voting is now open for the overall suitability for endorsement for Measure 2988. Option Number 1 is yes. Option Number 2 is no.

Looking for one more vote.

All votes are in and voting is now closed. The vote for overall suitability for recommendation for endorsement for Measure 2988 is 89 percent yes, 11 percent no.

CO-CHAIR SEPTIMUS: Okay. Well, I want to thank the developers for being nimble and --

DR. NISHIMI: Thank you for your patience.

CO-CHAIR SEPTIMUS: -- getting people -- any -- who is on the phone right now?

CO-CHAIR THRAEN: He had to drop off.

CO-CHAIR SEPTIMUS: Is that the only one that's on the phone? Okay.

Well, thank you very much for getting him on the phone.

DR. NISHIMI: Thank you very much. We

appreciate it.

CO-CHAIR SEPTIMUS: I think that helped a lot.

MEMBER MCGIFFERT: And can we say that we would really like to see your review measure being developed a little bit more and seeing that combined in the future with something like this because we really need it so --

DR. NISHIMI: Yes. We appreciate that.

CO-CHAIR THRAEN: Good luck with that.

DR. NISHIMI: It's a tough one, though, as you can imagine.

CO-CHAIR SEPTIMUS: Well, thank you very much.

So now we're five minutes ahead of schedule as we begin our afternoon. And so we have three measures we'll consider before the break that are PACE related: acquired pressure ulcers; fall rates; and fall rates with injury. So we know we had some folks who had to recuse themselves from this discussion.

(Off microphone comments.)

CO-CHAIR SEPTIMUS: Yes. I know. I'm getting there. I'm just introducing the afternoon session.

She's a great co-chair, by the way. She really keeps me straight and narrow.

Okay. So our developers here, are they -- are they here in person? Excellent. Excellent.

And then, I guess, Chris, you're going to --

CO-CHAIR THRAEN: No. It's Susan.

CO-CHAIR SEPTIMUS: Susan. I'm sorry. I'm sorry. Forgive me.

CO-CHAIR THRAEN: Susan -- she's here.

CO-CHAIR SEPTIMUS: Thank you, Susan. Sorry.

MR. LYZENGA: And, actually, Susan, could we ask you to quick introduce yourself and do a disclosure?

CO-CHAIR SEPTIMUS: Didn't see her. Sorry.

MEMBER MOFFATT-BRUCE: Good afternoon. It is afternoon, yes? I'm Susan Moffatt-Bruce. I'm Professor of Surgery and Biomedical Informatics at the Ohio State University Wexner Medical Center. I'm also the Chief Quality and Patient Safety Officer for our health system. I have no disclosures.

CO-CHAIR SEPTIMUS: So another Buckeye.

MEMBER MOFFATT-BRUCE: I am. I am.

CO-CHAIR SEPTIMUS: That's right.

MEMBER MOFFATT-BRUCE: So --

CO-CHAIR THRAEN: Susan, before you start, we ask the measure's developers to do a brief presentation and then --

MEMBER MOFFATT-BRUCE: Sure.

CO-CHAIR THRAEN: -- you can go from there.

MEMBER MOFFATT-BRUCE: Absolutely.

CO-CHAIR SEPTIMUS: You want to introduce yourselves, please?

MR. STEWART: Good afternoon. I'm

Mark Stewart. I'm the measurement and improvement lead at Econometrica. We have a measurement instrument development and support or "MIDS" umbrella contract for the Centers for Medicare and Medicaid Services or CMS. The measures that are being developed for PACE are under the MIDS umbrella.

With me is Dr. Nancy Dunton from the University of Kansas School of Nursing. She has experience with multiple quality measure sets, including the original national database of nursing quality indicators, or NDNQI, which were developed by the American Nurses Association. And joining us by phone is Ms. Tamika Gladney, from CMS. CMS is the steward for these measures that will be discussed today.

I thought it might be helpful to give a really brief background on the PACE programs that may not be known well nationwide. It's the programs for all-inclusive care of the elderly or PACE. This is a unique Medicare and Medicaid program with capitated funding administered by CMS

and the states. There are currently PACE organizations in 32 states.

PACE has a unique goal of keeping frail elders in the community and out of nursing homes. The population is relatively homogenous. They must be age 55 or older. They must have Medicare or Medicaid greater than 90 percent or dual eligible and be certified by the State as nursing home eligible.

Each PACE participant is living in the community with a designated caregiver and they provide truly interdisciplinary care. The care team consists of physicians, nurses, therapists, social workers, dieticians, personal care aides, transportation drivers, and others. The care and services include: clinical care; physical and occupational therapy; personal care; transportation, including to specialty appointments; recreation; socialization; and meals are provided at Adult Day Center. And the centers also provide adult day care and make modifications that may be necessary in the

participants' homes.

MEMBER SMIRZ: PACE -- is this live or -- sorry. Thank you. There's been some research on PACE programs -- evaluation studies in the past that show that they save money, extend participants' lives and reduce time spent in congregate care, hospitals and rehabilitation facilities.

The evaluations suggest that these reductions in utilization come from reduced length of stay in these settings rather than lower rates of entry. Other outcomes compare favorably to other programs for frail elderly. They maintain functional status, improved instrumental activities of daily living, and lower cost than nursing homes.

PACE programs must complete quality assessment performance improvement projects, although, historically, the quality measures have been subject to rather continual change in reporting to CMS. This is not consistent with having standard quality and safety measures that

can be compared across sites and over time.

CMS has initiated a process for developing standardized quality and safety measures for PACE, as Mark mentioned. The submitted measures are harmonized with existing NQF-endorsed measures for falls and falls with injury and are harmonized with a previously-endorsed measure for pressure ulcers, all three of which are in hospital settings and nursing home settings, were primarily harmonized with the hospital settings.

All address important outcomes in the frail elderly population and our analysis has demonstrated that they are reliable and valid for use in these programs. And CMS is considering the use of these measures in accountability applications within the next two years.

CO-CHAIR SEPTIMUS: So frail and elderly is now over age 55?

(Laughter.)

CO-CHAIR SEPTIMUS: Almost there? They're raising the bar, folks. We're lowering

the bar.

(Laughter.)

CO-CHAIR SEPTIMUS: All right. So we're going to -- anyone have any questions about that intro before we go measure-by-measure?

Great intro. Appreciate that very much.

So the first one is Measure 3000, PAGE-Acquired Pressure Ulcer Injury Prevalence Rate and, as we do -- you'll discuss the specs and validation of your measure. And then we have Susan's going to discuss it step-by-step for endorsement.

MEMBER SMIRZ: Okay. Sorry. The specifications of the measure or the first measure being the participant fall rate. You will hear me say "participant" rather than "patient" throughout this because I've been --

MR. LYZENGA: This is Pressure Ulcer 3000.

MEMBER SMIRZ: Oh, pressure ulcers -- sorry. Pressure ulcers. Yes.

The pressure ulcer rate is a prevalence rate. It's the number of PACE participants that have one or more pressure ulcers in -- during the quarter expressed as a percentage of all PACE participants during the quarter and have been on the PACE enrollment registry for at least one day out of the quarter.

The -- how deep do you want me to go into this?

CO-CHAIR SEPTIMUS: As far as you want.

MEMBER SMIRZ: Oh, boy. Okay. So the --

CO-CHAIR SEPTIMUS: Convince us we should endorse it.

MEMBER SMIRZ: Okay. All right. So, as I said, this measure is harmonized with previous pressure ulcer measures. It's a prevalence rate so it reflects the care -- the burden of care for PACE programs. It is -- has a number of admission and exclusion criteria, given that definition.

PACE participants are included if they've been on the PACE enrollment list for at

least one day, they have -- they're included -- the pressure ulcers are included if they were not present on enrollment -- I mean they're excluded if they were not present on enrollment. And they're excluded if they were acquired during a hospital stay or a nursing home stay because, technically, PACE is responsible for those participants and pays for their care no matter where they are but they're not -- their program is not actually -- the care that resulted in the pressure ulcers occurred outside their home or assisted living home -- usual home -- place of care.

So the -- if they come out of care in a congregate setting, if the pressure ulcer appears less than 24 hours after they return home it's also excluded because it was possibly then required in the congregate care setting.

Reliability testing -- we -- well, validity testing. We used a sample of both PACE experts who are knowledgeable about the population and the program and academic-type-measurement people who are specialized in pressure ulcers. We

had a list of a number of -- like a dozen academic experts.

So the explanation -- the -- they were given instructions that were -- and definitions -- instructions for care -- data collection for review and asked to comment on each of the elements, whether they thought that they were valid or not, both the numerator, the denominator, the exclusion and inclusion criteria as well as the process of data collection.

And they had -- they -- the statistics were good in terms of percent agreement or the ICBI measures at -- I don't have it in front of me -- seventy-some percent. Okay. For -- the ICBIs, overall, were .75, which is moderate. The numerator and denominator and rate as a whole were .88, so very good -- high. The evidence that it distinguishes good care from poor care was moderate because of the frail nature of this population.

Exclusions and -- for the numerator and denominator ranged from 1 to as low as .88, so very high. Exclusions for the numerator rated from --

ranged from 1 to .75. There was some discussion among the comments provided by the validity experts or the experts that we used about whether or not we should be including what is known in the field as Kennedy terminal ulcers, although that's not a stage and it's not necessarily a pressure ulcer, and whether or not we ought to be counting other kinds of skin breakdown, which we reject both of those things because we are focused on pressure ulcers specifically because you can only -- you improve care for pressure ulcers one way, while venous ulcers or diabetic ulcers or other kinds of skin breakdown are handled differently. So we need to be clear about what the measure is so that quality improvement can occur.

The reliability -- the validity study was done on data collected from January and February of 2015. It -- we had PACE organizations select a -- some of them have one site. Some of them have multiple sites. We asked them that they collect data on every PACE participant for those two months in their oldest site and then we used

signal-to-noise analysis to analyze the data.

And the signal-to-noise coefficients were -- let me see --

CO-CHAIR THRAEN: .78?

MEMBER SMIRZ: Yes. The signal-to-noise coefficients were --

CO-CHAIR THRAEN: .73 and .7 -- .83.

CO-CHAIR SEPTIMUS: Point seven something, right?

MEMBER SMIRZ: Yes.

CO-CHAIR THRAEN: .73.

MEMBER SMIRZ: So they're moderate.

CO-CHAIR THRAEN: And .83 for greater than stage three?

MEMBER SMIRZ: For greater than two, yes.

The -- we tested the -- we looked at terms for risk assessment and for risk adjustment in a number of ways but -- and, particularly, we tested whether or not we should adjust for age and gender or age and sex and found that there were no significant correlations of the rate with those

demographic variables. So the reliability studies indicate moderate reliability.

The --

CO-CHAIR THRAEN: Can I give you a break a minute --

MEMBER SMIRZ: Okay.

CO-CHAIR THRAEN: -- and see -- are there any questions so far for anything that she's covered?

MEMBER MOFFATT-BRUCE: So I just -- I do have a question. When we were talking about the numerator, you -- this is inclusive of all pressure ulcers of all stages, correct?

MEMBER SMIRZ: Correct.

MEMBER MOFFATT-BRUCE: Okay.

MS. HAMMERSMITH: But we collect it by stage and we did -- in the -- in the reliability study, we collected it by stage and defined stages as you have in your documentation so that CMS could then decide if they wanted to use for accountability any particular set of stages.

MEMBER MOFFATT-BRUCE: Sure. And

then, to that end then, when I look at your reliability data, it looked like the majority of them were unknown stage.

MEMBER SMIRZ: Yes. I think that there were -- there were a number of things we tested about pressure ulcers, including risk assessments and prevention activities as well as things like stage and it -- they were not in the records that they had access to, so they were not recorded in the -- in many cases. Age and gender or sex were recorded more frequently than risk assessment or prevention activities. So those were deemed not feasible, basically.

MEMBER MOFFATT-BRUCE: Right. And we'll probably discuss that a little bit in the feasibility aspect of this?

MEMBER SMIRZ: Right.

MR. LYZENGA: And maybe we could get into the discussion so we can kind of walk through the criteria. Let's start on evidence and --

CO-CHAIR SEPTIMUS: It's a great intro. Most of the other measures we've

considered, I believe, on decubitus ulcers to date have looked at Grade 3 and Grade 4, which --

MEMBER MOFFATT-BRUCE: Correct.

CO-CHAIR SEPTIMUS: -- is maybe why Susan asked that question.

MEMBER MOFFATT-BRUCE: Yes.

CO-CHAIR SEPTIMUS: So just to let everybody know, that's a little bit different. So Susan's going to lead us through our discussion about --

MEMBER SMIRZ: Okay.

CO-CHAIR SEPTIMUS: -- evidence, et cetera.

MEMBER SMIRZ: Okay.

CO-CHAIR SEPTIMUS: And, as we go, we may have some additional questions for you. So --

MEMBER SMIRZ: All right.

CO-CHAIR SEPTIMUS: So thank you very much for the intro.

MEMBER SMIRZ: Yes.

CO-CHAIR SEPTIMUS: Susan, let's start off with the evidence.

MEMBER MOFFATT-BRUCE: Absolutely and I will do my very best to do this justice.

So, first of all, this is a new measure. The evidence is that pressure ulcers do cause not only physical but psychological harm and really should be measures in all settings. So I do think that the evidence that they've documented here is that they are present, the are present in all kinds of care settings.

And at this PACE -- and I thank you for the summary, because I had to look up what PACE was and do my own inquiry -- should be included in that as it's a very important population and vulnerable population that we serve.

So I have no additional comments on the evidence. Relative to --

CO-CHAIR THRAEN: No. Wait a minute.

Sorry -- I'm sorry. I apologize.

CO-CHAIR SEPTIMUS: Any -- yes. Yanling?

MEMBER YU: Yes. Same if you look it up or what they say it is.

MEMBER MOFFATT-BRUCE: Yes. Yes.

MEMBER YU: It said the service deliver setting include home. So --

MEMBER MOFFATT-BRUCE: It's mostly home, isn't it?

MEMBER YU: So my question is -- I'm just -- a matter of educating myself. How does pressure ulcer would be documented or determined or discovered in home setting?

MEMBER SMIRZ: The PACE Program -- this is one factor that we failed to mention. The PACE Program has a care provider in every home. It could be a relative or it could be somebody that's hired for providing care but they have daily care by someone who documents problems, who assists them with activities in their homes.

MEMBER YU: Okay.

MEMBER SMIRZ: And they are assessed periodically by nurses who visit the home.

MEMBER YU: Okay. And then there's one -- under the evidence it said a pressure injury incident rate for a pays-per-one are not available.

But the evidence quoted other statistical observations from different settings. One is 0 percent to 70 percent in home care setting. I'm just wondering is this the numbers that you think is appropriate that we use for this as evidence?

MEMBER SMIRZ: I think it's the -- it's the closest number that we have because the people at PACE are actually more frail than people in home care because they are nursing home eligible --

MEMBER YU: Okay.

MEMBER SMIRZ: -- and would otherwise be placed in a nursing home if it were not for the PACE program.

MEMBER YU: Okay. All right. Thank you.

CO-CHAIR SEPTIMUS: Okay. Any questions specifically about the evidence?

Yes, Pat?

MR. LYZENGA: And just a quick clarification. This is an outcome measure. It's the first outcome I think we've looked at today. So the question on evidence here is whether there's

a rationale connecting at least one process or intervention that a measured entity can do to impact the outcome.

MEMBER QUIGLEY: Thank you. Pat Quigley. And my question in relationship to the evidence for this patient population and the PACE Program, because these are old people who have frailty and meet the criteria for nursing home admission is --

CO-CHAIR SEPTIMUS: These are senior citizens, please.

MEMBER QUIGLEY: I am one of those. I meet this --

CO-CHAIR THRAEN: A little bit of clarification. It's not limited to senior citizens.

MEMBER SMIRZ: Anybody over 50.

MEMBER QUIGLEY: It's 55 and older.

MEMBER SMIRZ: 55 and over.

MEMBER QUIGLEY: 55 and older but they meet frailty criteria and the -- and are eligible for nursing home admission. Yes.

CO-CHAIR THRAEN: Then disabled.

MEMBER QUIGLEY: Yes.

MEMBER SMIRZ: They have to be eligible, though, for nursing home.

MEMBER QUIGLEY: Yes. So, for this indicator on pressure ulcer prevalence is how is this -- how do we determine the structure and process to prevent new pressure ulcers? This is a prevalence measure versus an incidence measure.

And you had mentioned, Dr. Dunton -- and thank you so much for that overview -- that pressure ulcer prevalence upon admission to the PACE Program is excluded. That wasn't in all of our discussion present on admission. But, in this population, there are many who will develop pressure ulcers that are absolutely preventable because they are frail old people who don't have the healthy tissue or the abilities to be able to not prevent a pressure ulcer.

So where is the evidence to support the structure and process for this population for a prevalence study -- a prevalence measure versus an

incidence measure?

MEMBER SMIRZ: Oh, thank you. The -- we did not collect any data on the structure and process measures. However, the caregiver -- once a -- once a pressure ulcer's identified or somebody's identified at risk of a pressure ulcer on their periodic evaluation by a physician or visiting nurse, will receive care in the home that is appropriate.

And, for people who are bedridden, of course, that's the usual turning, moisture management -- moisture management, whether they're in bed or up, as well as nutritional support. I think nutritional support and pressure-reducing surfaces -- all of those measures -- all of those prevention measures can be employed in the home by the caregiver and by nurses who will come by to visit the participant.

Structure is -- there is -- there's no variation in structure, really, I think because there is a caregiver in the home. So it's one-on-one.

MEMBER QUIGLEY: Thank you so much. But my question was really more relevant to having a prevalence measure versus an instance measure.

MEMBER MOFFATT-BRUCE: Right.

MEMBER QUIGLEY: For instance, and this is done quarterly and for a very, very frail, debilitative patient population. So that's where I really question the evidence that was presented and even using NDNQI.

Because NDNQI data or that model of care is very different than those who are living in the home versus those who are living in the assisted living. They're very different contexts, very different settings of care.

So I just question the structure and the process to be able to support this outcome.

CO-CHAIR SEPTIMUS: Thank you, Pat. Any other? Seeing none, we will vote on the evidence.

MS. QUINNONEZ: We are now voting on Measure 3000, PACE measure. It's acquired pressure ulcer injury prevalence rate. We are

voting on the evidence of this measure. Option Number 1 is yes. Option Number 2 is no. Option 1 is yes. Option 2 is no.

Okay, thank you again. All right, we are re-voting on Measure 3000, on evidence. Option Number 1 is yes. Option Number 2 is no. Actually, yes.

Okay. All votes are in. This voting is now closed. The vote for evidence of Measure 3000 is 89 percent yes, 11 percent no.

CO-CHAIR SEPTIMUS: Okay, Susan, let's to Gap.

MEMBER MOFFATT-BRUCE: Okay. So to go to Gap. So the Gap around this particular measure was calculated -- and I just want to make sure I get this clear -- from a sample of 50 sites out of 114 potential sites. Yet, only a total of 29 of the sites submitted data.

Having said that, the rate then became -- or the mean was 0.81 per 100 participants for Stage 3 and above and, I guess, 1.85 for every 100 participants for all types of pressure ulcers.

The inclusion criteria for these patients are by anybody living at home or in assisted living facilities. And then the exclusions are very clear.

With that performance being measured and this inclusion and exclusion criteria which are fairly well delineated, I think that the developers have demonstrated that there is a performance --

CO-CHAIR SEPTIMUS: Susan, let me just talk, I'm sorry, for a second. You're, I think, going one step ahead.

MEMBER MOFFATT-BRUCE: Okay, very good.

CO-CHAIR SEPTIMUS: So I apologize.

MEMBER MOFFATT-BRUCE: That's all right.

CO-CHAIR SEPTIMUS: So what we want to do is the performance gap.

MEMBER MOFFATT-BRUCE: Okay.

CO-CHAIR SEPTIMUS: Is there a performance gap?

MEMBER MOFFATT-BRUCE: Okay. The --

or the data here would demonstrate that there is a performance gap.

CO-CHAIR SEPTIMUS: Here's my question to all of us. I think they are using other settings as being illustrative of a performance gap. But I don't think they have any prior information around the PACE program. Am I reading --

MEMBER MOFFATT-BRUCE: Well --

CO-CHAIR SEPTIMUS: Is my reading incorrect? I'm asking.

MEMBER MOFFATT-BRUCE: No, I -- my -- maybe the developers can explain this, but my understanding is that they actually took it from PACE sites, this data.

These 29 -- so they asked 50 sites to submit data. They got 29 sites to respond. Am I not reading that correctly on --

CO-CHAIR SEPTIMUS: No, I'm asking. I mean, there was a little bit -- that was a little confusing to me because, as I'm reading the document here, it says here that, "strong evidence for a pressure ulcer is highly impactful,

preventable -- oh, I'm sorry.

"There's solid evidence performance gaps in variation in care exists in other healthcare sites such as Acute Care, Long-term Care and Home Care setting. However, there is no current evidence that it exists in the PACE program per se."

MEMBER MOFFATT-BRUCE: Right. And then when they come down into the next paragraph or the next setting, they speak to --

CO-CHAIR SEPTIMUS: Right.

MEMBER MOFFATT-BRUCE: -- having collected this data. So it is a bit contradictory, on the same page.

CO-CHAIR THRAEN: The gap, the data that is under performance --

MEMBER MOFFATT-BRUCE: Performance Gap 1 --

CO-CHAIR SEPTIMUS: They --

CO-CHAIR THRAEN: Performance gap data is really not looking at performance gap. It looks like it just measuring --

MEMBER MOFFATT-BRUCE: Incidence.

CO-CHAIR THRAEN: -- the rate --

MEMBER MOFFATT-BRUCE: Incidence.

CO-CHAIR THRAEN: -- of incidence or rate of prevalence in the setting. So it's really not addressing the performance gap.

MEMBER MOFFATT-BRUCE: Right.

CO-CHAIR THRAEN: Is that correct? People agree with that or not?

MEMBER MOFFATT-BRUCE: I think that's the only conclusion.

MEMBER SMIRZ: May I speak?

MEMBER MOFFATT-BRUCE: Yes.

CO-CHAIR SEPTIMUS: Please.

MEMBER SMIRZ: The apparent conflict is that, prior to this study, there were no measures of pressure ulcers in the PACE population.

But then we did collect data on the PACE population and provided the statistics that you referenced. From my usual way of thinking about this, was that, to show a gap, that you would look at the range of rates.

MEMBER MOFFATT-BRUCE: Right, variation. All right.

MEMBER SMIRZ: Yes.

MEMBER WU: So I'm still, I mean, if you could explain to us what you think the gap was. Is that, is anything other than zero a gap?

And the fact that there were no significant differences between the sites, it would seem to me that variation in, you know, in outcome would, for me, be a generally be a demonstration that you could attain a higher rate or a better rate in one setting or another.

But the fact that there was not a significant difference, to me, does not, at least not on that criterion, support there being a gap.

MEMBER SMIRZ: Looking for the range. The --

MEMBER MOFFATT-BRUCE: Want to just read that, right there?

MEMBER SMIRZ: The range was from --

MEMBER MOFFATT-BRUCE: Was varied. Zero to 0.7 percent.

MEMBER SMIRZ: -- 0.31 to 5.60. So it was a fairly large range. And, recall that these are percentages.

MEMBER WU: Statistical testing is not everything. But it depends whether or not -- so P-values are not everything, but these differences were not statistically different from each other?

MEMBER SMIRZ: I don't think we test -- I don't think we calculated statistical tests, which was -- the number of reporting sites would have been perhaps interpreted with too much assurance.

CO-CHAIR SEPTIMUS: Tell me if I'm reading this wrong because it said the 29 PACE sites that were just --

MEMBER MOFFATT-BRUCE: Surveyed.

CO-CHAIR SEPTIMUS: -- referenced to were looking at Stage 3 or higher. And I think that you said --

MEMBER SMIRZ: Oh, Stage 3 or higher has a -- that was all. Stage 3 or higher had --

MEMBER MOFFATT-BRUCE: Point eight one.

MEMBER SMIRZ: -- a range of 38 -- .38 percent to -- or zero percent to 3.47 percent.

CO-CHAIR SEPTIMUS: No, but that -- sorry, wasn't the measure that you're presenting any decubitus ulcer, not 3 or 4?

MEMBER SMIRZ: Yes.

CO-CHAIR SEPTIMUS: Okay. That's why I'm -- I was a little confused. So maybe other people are not, so maybe they can help.

MEMBER WU: Just to clarify, I think my reading of what's up here is that, so that, those P-values down there are the comparison between those who are affiliated with academic medical center, yes/no, and metropolitan versus micropolitan.

So those are not significant. But if you look above, it looks like there is variation where it says the number of participants with PACE acquired pressure ulcers, for every 100 participants and equals 28, there's 1.85.

And I guess one of the questions, maybe, for the developer is that there, it looks like there

are two, four numbers after that. Are those the confidence interval? What do those represent?

MEMBER SMIRZ: The four numbers are mean standard deviation, median, minimum, maximum. The -- I'm sorry, the formatting was lost in the form.

CO-CHAIR SEPTIMUS: So, Jesse, interpret that for us, please.

DR. PINES: So it looks like there was no statistical tests done to assess whether there was variation across the sites. But if you look at the min and max, it looks like there is variation there. That's my interpretation.

MEMBER MOFFATT-BRUCE: So variation doesn't --

DR. PINES: That's the -- that would be the min and max with the average of 1.85.

CO-CHAIR SEPTIMUS: Leslie?

MEMBER SCHULTZ: Maybe I'm being simplistic. It says that incident rates are not available in this population. It's a new measure. You don't have any historical information on it.

But then you give the expected ranges for pressure ulcers in nursing home and rates for persons receiving home care. And there's ranges here, although it's a little dated, 2001.

So I think it would be interesting to see, you know, do these proportions look anything like those other care settings? And then at the different sites, is -- you know, there's some really bad sites and some really good sites and probably --

MEMBER SMIRZ: Yes, I'm sure there are good and bad sites. The expected level of pressure ulcers would be higher in PACE settings, because of their frailness than in home care.

It might be lower than those hospitalized and those in nursing homes. So, but we didn't have the same data on all of those for some recent.

Home care data used to be available from OASIS on the Web but had been removed from the website. So we did not have that. But, and so what we provided, then, is to say, historically,

these have been presented in other places that have shown variation, and there is variation in the PACE sites as well.

CO-CHAIR SEPTIMUS: Pat, did you have another question as to --

MEMBER QUIGLEY: Yes, thank you. And my question, in relationship to gap, besides those studies that was done here is, in relationship to the quality management program that exists for PACE already.

Acknowledging that PACE has been around since 1997, historically, PACE does not have any data on pressure ulcers in this patient population?

MEMBER SMIRZ: Not that's been publicly reported. CMS has had improvement projects for, if -- I hope I'm presenting this correctly. CMS has had improvement projects that have focused on different aspects of care each year.

And so the PACE sites that volunteer for that program have reported to CMS on the results on pressure ulcers previously, but those data were

not available to us.

MEMBER QUIGLEY: And so, in relationship to the gap -- and thank you for that response -- in relationship to the gap in terms of structure or process, PACE is a program that manages care.

It is not a provider of care other than going in and completing an assessment or monitoring and coordinating care. So --

MEMBER SMIRZ: They do care coordination but they also have physicians --

MEMBER QUIGLEY: Right, but they don't actually provide care.

MEMBER SMIRZ: Oh, yes, they do.

MEMBER QUIGLEY: Oh, like the skin care management? Because there are patients in PACE who can have -- this is my question -- that also has, can it receive home health care?

MEMBER SMIRZ: Yes.

MEMBER QUIGLEY: Yes. So in home health care, CMS has the OASIS program. And the OASIS program, for home health care, monitors

pressure ulcers, just like FALLS. So does PACE not have data from those who are in home health care?

MEMBER SMIRZ: I do not know if we could have -- to my knowledge, the data from OASIS that would apply to PACE settings, that subset, was not available to us, nor was the data that CMS has collected from volunteers which would probably be via a sample of those that have more active quality improvement programs or have the staffing or the history to do that kind of work.

Also, were not available, but those data wouldn't have been represented even, in any way.

MEMBER QUIGLEY: Thank you.

CO-CHAIR SEPTIMUS: Albert, did you have -- okay, I don't see other --

MEMBER MOFFATT-BRUCE: Jason?

CO-CHAIR SEPTIMUS: Jason, are you there?

MEMBER ADELMAN: What section?

CO-CHAIR SEPTIMUS: Nice to hear you.

MEMBER MOFFATT-BRUCE: Performance

gap.

CO-CHAIR SEPTIMUS: I'm sorry, what was the question, Jason?

MEMBER ADELMAN: No, I had one question. So it's something earlier. I'm sorry, I'm embarrassed to ask the question, but could you please just remind us, what does it mean when they said that this measure's harmonized with existing measures that are already endorsed, like the NDNQI pressure ulcer measure?

What does that mean, again, to be harmonized with?

MR. LYZENGA: I guess it can mean different things in different instances. Often what you look at is, for example, are definitions aligned, harmonized? I think there can be different dimensions of harmonization between similar measures.

I would ask the developers, you know, that the --

MEMBER SMIRZ: Proposed measures -- this proposed measure and the NDNQI measure and the

former NQF endorsed measure use definitions of a pressure ulcer and the stages of a pressure ulcer as specified by the National Pressure Ulcer Advisory Panel, NPUAP, which is the national body that defines how pressure ulcers are identified and --

MR. LYZENGA: I just -- we use the word harmonized sometimes, even outside this measure. And I just want, I'm trying to understand the value of this on top of NDNQI and does harmonized mean different but synchronous?

Or like, because if it was, if it overlapped too much, then we would have an issue with it. So -- and that's what we're stating that this is, that it is different.

Because I see lots of overlap. But I'll just --

MEMBER SMIRZ: Yes, there is lots of overlap. The thing that's different is that there are different exclusion criteria for PACE programs, because it's a different setting, to go along with what's the responsibility of PACE

programs as opposed to they're in the hospital, so we know the hospital's responsible for the care of hospital-acquired pressure ulcers

So this is PACE-acquired pressure ulcers but -- and the definition of a pressure ulcer's the same but the exclusion criteria, basically for the denominator, are different.

MEMBER ADELMAN: Well I have some issues with the validity and --

MEMBER SMIRZ: But that --

MEMBER ADELMAN: But I'll wait until we get to that subject.

CO-CHAIR SEPTIMUS: Yes, right.

MEMBER ADELMAN: And we can, we'll have a discussion around related and competing measures for this --

MEMBER SMIRZ: Right.

MEMBER ADELMAN: -- later on. And we can talk about that. Generally, harmonization applies to measures that are related but aren't deemed to be close enough to be competing with each other.

And those ones you would want to make a decision about best in class, if they're related, you see a justification for having both, then you might want to approach the developers and say, can you harmonize things like definitions, for example, to reduce the burden on those collecting data.

But, again, we can talk about that at the related and competing.

CO-CHAIR SEPTIMUS: So, Lisa, one more comment and then I think we probably ought to vote on the gap. Oh, I'm sorry -- and Iona. And then we'll vote on the gap.

MEMBER MCGIFFERT: Okay, I just need to figure this out a little bit. When the caregiver is the person that's managing the care for this person on PACE, how does that, let's say it's their brother or some, you know, how does that person document that the PACE participant has a pressure ulcer?

MS. GLADNEY: Hello? But, yes, this is Tamika Gladney from CMS. Hello, everybody.

CO-CHAIR SEPTIMUS: Hello.

MS. GLADNEY: Hello. I just want to give a little bit of clarity about what the young lady just asked.

So in the PACE program they are provided with an IDT. That's the Interdisciplinary Team. It consists of physician, physical therapists, a nurse, social worker, occupational therapy, transportation. So all of it, it's about 11 services that they get just for that one participant.

The IDT takes care of the whole entire PACE organization and/or participants. But those services are then provided to each one instead of them being in a nursing home and/or being in a hospital receiving these services.

So for a participant who is at home, the IDT do, they complete an assessment and, along with their physician on this IDT. And they say, okay, this particular participant, they can't move around as much. We think that they're going to need home care two times a week, at least.

And in addition to that, with the home care, they provide what type of home care. Are they going to need, you know, that? They're going to need laundry. Are they going to need their food made for them?

Even though they may have a caregiver at home, they also would have a nurse that would come out and do those kind of physical assessments at home.

That information then is documented and brought back to the IDT meeting where all the other 11 disciplines who take care of that participant can hear the information. So that's kind of how the information is identified for these participants.

MEMBER MCGIFFERT: So that means that the information, the IDT, the team, decides this person --

MEMBER QUIGLEY: Has a pressure ulcer.

MEMBER MCGIFFERT: -- did get the pressure ulcer?

MS. GLADNEY: That is correct.

CO-CHAIR SEPTIMUS: Yes, they have to put you --

MS. GLADNEY: So that team decides how often it comes, they come out. They, if they're coming to the center versus getting home care, they can actually get both. If they're coming to the center then they also, you know, review the participant's integumentary system, et cetera, needs at home, et cetera.

CO-CHAIR SEPTIMUS: Yes.

MS. GLADNEY: Their, also, family gets involved as their advocate or, you know, significant other, be also brought into the care plan on a regular basis.

So they, you know --

MEMBER QUIGLEY: And this is Pat's voice, Pat Quigley's voice. But you can also bring in home health care, a home health care agency. The IDT team can decide to bring in a home health care agency.

CO-CHAIR SEPTIMUS: I hate --

MS. GLADNEY: But part of it, part of

the plan care net?

MEMBER QUIGLEY: Yes.

CO-CHAIR SEPTIMUS: Hate to cut this discussion off, but I'd really like to just confine comments to finish the gap analysis discussion. And then we'll move to all the other -- I hate to cut great discussion off, but I just want to get passed the gap analysis.

So, Iona, if you have that one, one more comment.

CO-CHAIR THRAEN: So just two observations. Well, one question, how often is the assessment done? This is to the CMS representative. How often is this assessment done?

MS. GLADNEY: It's based on individualized needs of the participant.

CO-CHAIR THRAEN: Okay. All right. So how I've seen it operationalized in Utah is the, it's really intended for the younger disabled population so they don't have to live in a nursing home.

They have to go into the nursing home for X period of time. They're assessed that they could live independently, it cut costs on the Medicaid side to be able to support them in the home. And that's how I've seen it used primarily.

And the second piece I wanted to say is this is sort of akin to the dialysis centers. This is a population that hasn't been monitored that closely. And so this is an opportunity similar to dialysis to kind of get them into that loop of looking at this issue.

I don't know what the rates, how this compared to the rates of the skilled nursing versus the hospitals versus home health, but it's a high risk disabled population, I would say, is the best way to describe it, I think.

You know, your MS patients, your Lou Gehrig's patients who do not want to live in a nursing home who have some social support at home. And this helps them stay at home.

CO-CHAIR SEPTIMUS: Okay. Around the gap?

MEMBER WANG: Just, if I could --

CO-CHAIR SEPTIMUS: Just about the gap?

MEMBER WANG: Yes.

CO-CHAIR SEPTIMUS: Okay.

MEMBER WANG: Yes, the clarifying question, so if a pressure ulcer is developed at home from the home setting, that's captured into this prevalence rate?

MEMBER SMIRZ: Yes. Yes, it is captured. And it would be in the clinical record maintained by the Interprofessional Team.

MEMBER WANG: And I have a follow-up. So I guess my question is, so because the home, the family member is being included into the care of this member and the measure is at the PACE level, are we kind of -- you know, so pressure ulcers can be developed if, let's say, the family doesn't take good care of the member.

So in my, in a way, are we deeming the PACE organization for the, for a high prevalence rate that is not entirely within their scope?

CO-CHAIR SEPTIMUS: Thank you. And, hate to cut off, is this about the gap?

MEMBER SMIRZ: Yes.

CO-CHAIR SEPTIMUS: Okay, go ahead.

MEMBER SMIRZ: Yes, the mechanics, when the mechanics when the family member is involved -- I've been involved with that.

And basically it's not that I'm going to make the diagnosis but I am told what to do -- and this is part of the linkage that's going to come into the measure and how it's documented.

So I see something. I make a phone call. Somebody with clinical expertise comes and looks. And then they make the diagnosis. So that's how the family fits in with picking up the prevalence.

CO-CHAIR SEPTIMUS: Yes, speaker respond.

MR. STEWART: Yes, that's true. And one point we left out, these PACE organizations maintain clinics under the same roof as the adult day center. And PACE participants spend, on

average, two days a week in the day center and the clinic where they're seen by physicians and nurses.

CO-CHAIR SEPTIMUS: Okay, let's -- Chris, you're ruining my schedule here.

MEMBER COOK: I have to ask this from a point where I'm conflicted. And I have to give commendation to CMS because they can do this without coming to NQF and put this in and go ahead with it.

So I commend CMS for following through the process to come to a multi-stakeholder body, look at this from an evidence.

From our standpoint, as committee members, when we look at this and we're evaluating it as an outcome measure, there's very little information in this area.

This is a very important topic, very critical patients who absolutely need to have this done, but there's very little evidence. So as in information gap, are we to be doing this off we see this as reasonable clinicians, knowing that this would be a problem that causes major psychological

and physical damage to the patients if it's not being done correctly?

Or are we looking directly just at the evidence of what's there for pressure ulcers in in-home settings in the PACE program? So point of clarification for our esteemed leaders.

CO-CHAIR SEPTIMUS: So I guess what you're asking is, is the evidence strong enough in, specifically for PACE participants, that there is a gap or are we extrapolating from other settings to -- and knowing this is an important measure? Is that what you're asking, Chris? Okay.

CO-CHAIR THRAEN: I think the performance gap is that we don't know. That's the performance gap, is that this is a, again, going back to dialysis, this is the first step for trying to understand this frontier of care that we don't really understand what's going on in that environment.

It could be that their rates are just as comparable to any of the caregiving environments, the professional caregiving

environments versus the, you know, loved-one caregiving environments.

It could be that the gaps are that there may not be any performance gaps. But we have no clue. So this is sort of that first step.

MEMBER COOK: But this is an outcome measure. The previous measure, just to let you know, was a process measure.

DR. PINES: Just to --

CO-CHAIR SEPTIMUS: Go ahead, Jesse.

DR. PINES: Just to --

MEMBER QUIGLEY: I would like to say, Mr. Chairman, that -- and to Iona -- about --

CO-CHAIR SEPTIMUS: You can call me Ed.

MEMBER QUIGLEY: Well, I just respectfully disagree because the analysis studies that have been done by Mathematica, that Nancy started to allude to, they compare PACE with home care, that they were similar in their outcomes.

But with the issues in the prior research, which wasn't presented, is that the PACE program did not control for fidelity, the integrity

of the Interdisciplinary Team and had difficulty recruiting people into the Interdisciplinary Team because it's a very special population, people with frailty, ready for nursing home placement.

So the control for fidelity and then implementation of the model is what's not there. That's the structure and process piece which I thought outcome measures must have, that they have to be able to show this link between structure and process for the outcome.

But there has been a comparative analysis. It was published in 2008 for the PACE program.

DR. PINES: Yes, just to clarify what we're voting on here. So this is performance gap for this particular measure which the developer did present some data, which is up here, demonstrating that there is variation across the 28 sites.

And you can see the data up there. So with the, you know, you've got the mean, the standard deviation of 1.4 and you see the range there. So the question, specifically, here is

around is that a sufficient performance gap?

CO-CHAIR SEPTIMUS: Thank you. So with that, let's vote.

MS. QUINNONEZ: Voting is now open for performance gap for Measure 3000. Action Number 1 is high. Action Number 2 is moderate. Action number 3 is low. And Action Number 4, insufficient.

CO-CHAIR SEPTIMUS: This may have been the longest discussion on gap. What is this now, three years?

MS. QUINNONEZ: All votes are in. Voting is now closed. The vote on performance gap for Measure 3000, zero percent for high, 44 percent moderate, 17 percent low and 39 percent insufficient.

MR. LYZENGA: So we've got another gray zone situation here. Consensus not reached, I think. So, again, we'll move on to the next criterion and revisit this after the comment period.

CO-CHAIR SEPTIMUS: Susan --

MR. LYZENGA: We will not take an overall vote on endorsement for this one.

CO-CHAIR SEPTIMUS: Thanks. Susan, you're on for the next one here.

MEMBER MOFFATT-BRUCE: The next measure or the next part? Next part, okay. Okay. Ahead of myself, sorry.

So this is the reliability. Okay, very good. This is complicated. I'm just simple surgeon. I'm just telling you.

CO-CHAIR SEPTIMUS: I feel like we're talking about the visiting angels here.

MEMBER MOFFATT-BRUCE: Yes. Yes, yes. So reliability, so under reliability. So, right. That's not what I have. Okay.

CO-CHAIR SEPTIMUS: We're looking at the specifications.

MEMBER MOFFATT-BRUCE: Yes.

CO-CHAIR SEPTIMUS: The inclusion and --

MEMBER MOFFATT-BRUCE: Yes, that's what I thought, okay.

CO-CHAIR SEPTIMUS: -- external and economic data.

MEMBER MOFFATT-BRUCE: All right.

(Off microphone comments.)

MEMBER MOFFATT-BRUCE: Okay, so the inclusion criteria are those that are living at home or in an assisted living facility. They include all types of pressure ulcers and present on admission or the ones -- no?

Acquired elsewhere are excluded, including if these patients, I presume -- or these participants, I presume, if they went and had an in-patient stay, developed a pressure ulcer, that they would not be included in this.

The exclusions are fairly well delineated in that they have to be -- if they were not -- they have to be, have been in these, this program for at least one day out of the quarter.

And they are excluded if they are in a hospice facility or a nursing home facility, skilled nursing facility or a rehabilitation center.

When I look at their, my comment previously was around, obviously, these are all types of pressure ulcers including what they're categorizing as, obviously, unstageable, deep tissue injury and then this category of unknown.

And I have a question for the developer on that, how that actually, how these actually get categorized in that we have care providers that may not be trained in skin assessment of all different types of expertise that are -- I know in an acute care setting, documenting that it's a pressure ulcer's very challenging. I would imagine that this would be, similarly.

That was my biggest question for the developers, if they could kindly comment on that?

MEMBER SMIRZ: Unknown is not about the stage of the pressure ulcer. It's that the stage was not recorded in the clinical record.

MEMBER MOFFATT-BRUCE: Right. And so that, I think that comes back to my other concern, is how -- because you spoke that this data comes back to the multi-disciplinary team to get

validated and determined to truly be a pressure ulcer?

MEMBER SMIRZ: Yes. It's -- well, it starts with the team. Actually, if a caregiver says, I think there's a problem or they're at the PACE center and, three days a week, and have, you know, questions somebody about a sore spot, then they get an evaluation by someone qualified to identify and stage a pressure ulcer.

MEMBER MOFFATT-BRUCE: What does someone qualified mean?

MEMBER SMIRZ: It means they're a nurse or a doctor.

MEMBER MOFFATT-BRUCE: I don't think most doctors can stage ulcers very well. But, so would --

MEMBER SMIRZ: Let me just say that --

MEMBER MOFFATT-BRUCE: We don't. You know, physicians do not stage ulcers in our institution because nurses do it so much better.

MEMBER SMIRZ: And that could be what happens. We don't actually have information on

that process, specifically, with this.

Let me remind you that we've included Stage to give CMS the opportunity to restrict reporting to a certain set of stages. But we were collecting all stages. And so if physicians can't stage, then at least they can identify a pressure ulcer.

MEMBER MOFFATT-BRUCE: Well, okay.

MS. GLADNEY: This is Tamika Gladney from CMS. I would like to just make a note. Normally in our PACE organizations, the physicians do not do the staging.

We, the PACE organizations have CWOC nurses that are available. And those CWOC nurses with certain skills for certain wound identification skills then help educate the nurses that are on the team and, too, the family members.

CO-CHAIR SEPTIMUS: You want to go through the reliability testing then? I think the daily, everyone -- are the daily elements fairly well defined, Susan?

MEMBER MOFFATT-BRUCE: They are.

CO-CHAIR SEPTIMUS: Okay. What about the reliability testing?

MEMBER MOFFATT-BRUCE: The reliability testing that I see here in front of me reveals that they have, relative to all or those 3 and above, that it's 0.73 for all and 0.83 for Stage 3 and above for the signal-to-noise assessment, which would say that it's about moderate with a fairly large range, however.

The validity, yes. So the validity was done as previously described as well, using face validity only. And that was done by a national panel. And it does have a high ICVI indicating that it does have reasonable validity.

CO-CHAIR SEPTIMUS: Can I ask a question about that table, to the developers? Is, so we see Stage 3 and 4, the second line.

MEMBER MOFFATT-BRUCE: Yes.

CO-CHAIR SEPTIMUS: The first line is that --

MEMBER MOFFATT-BRUCE: Is all.

CO-CHAIR SEPTIMUS: -- all ulcers?

MEMBER SMIRZ: Correct.

MEMBER MOFFATT-BRUCE: It's all, yes.

CO-CHAIR SEPTIMUS: Okay. So obviously the reliability scores are higher with the Stage 3 --

MEMBER MOFFATT-BRUCE: Three and above.

CO-CHAIR SEPTIMUS: -- and 4 than 1 and 2. But it's, even with all stages, it looks like the reliability scores are in the moderate range, but clearly better with Stage 3 and Stage 4. Is that -- am I reading that correctly?

MEMBER SMIRZ: Yes.

CO-CHAIR SEPTIMUS: Okay.

MEMBER MOFFATT-BRUCE: And I presume, one more question, that the unstageable, the deep tissue and the unknown are included in the 3-plus? Or are they included in the all only? What makes up that 3-plus?

MEMBER SMIRZ: Right. I think it's, 3 and 4, you can make the presumption that deep tissue or unstageable are 3 and 4. But we did not

specifically include them in the 3 and 4.

MEMBER MOFFATT-BRUCE: Thank you.

CO-CHAIR SEPTIMUS: Okay, so I think the question for our committee really is are the methods and results of the validity testing -- I mean, reliability testing -- adequate? And is it sufficient to detect differences in performance? I think those are the questions that we need to answer.

MEMBER WU: Could I just get a clarification? Sorry.

CO-CHAIR SEPTIMUS: I haven't called on you yet, Albert. Albert?

MEMBER WU: Well, yes. Are you calling on me? Could you just clarify for me -- I don't understand exactly what test was done for reliability here.

Did you look at 28 individual ulcers and this was a test of whether or not that ulcer was a Grade 3 ulcer or not? Or what was the procedure that was done? It wasn't sort of a subject --

MEMBER SMIRZ: It was -- sorry. The

test that was done was done on 28 --

MEMBER MOFFATT-BRUCE: Sites.

MEMBER SMIRZ: -- sites.

MEMBER MOFFATT-BRUCE: 28 out of the 50 sites that were randomly picked out of 114 --

MEMBER SMIRZ: Correct.

MEMBER MOFFATT-BRUCE: -- potential?

MEMBER SMIRZ: Correct. At any rate, so it's all of the people with pressure ulcers in those 28 sites during two months. And the analysis was done with signal-to-noise, so looking at the, basically, the percent of variance within a site versus between sites.

The high, the difference between sites represents ability to detect differences in quality among organizations. And that's what the 0.73 shows as moderate.

MEMBER MOFFATT-BRUCE: And, but sites, you mean locations whereas not skin sites? You're -- I think that's where --

MEMBER SMIRZ: Correct, I mean location of the program, not a skin site. Yes.

MEMBER ADELMAN: Can you go to H and I where it says 2A-1 specification? It's the description of the inclusion and exclusion criteria.

MEMBER MOFFATT-BRUCE: It's up. It's --

MEMBER ADELMAN: It's higher up.

MEMBER MOFFATT-BRUCE: It's a little higher, yes. It's two more.

CO-CHAIR SEPTIMUS: For those who are looking, it's the 2A-1 reliability --

MEMBER MOFFATT-BRUCE: There.

CO-CHAIR SEPTIMUS: -- specification.

MEMBER ADELMAN: Those two inclusion criteria, are they -- is there an "and" between them or and "or" between them? Like, do you have to have the first bullet and the second bullet to be included? Or do you have to have the first bullet or the second bullet?

MEMBER SMIRZ: Or.

MEMBER ADELMAN: And then, so I get --

CO-CHAIR SEPTIMUS: Or.

MEMBER ADELMAN: I'm not -- can you just

scroll a little bit down, please. Stop. Sorry. I can't -- I just -- sorry. I can't reconcile -- go up a little bit. Sorry.

This sentence here, "Include participants where if something happens in 24 hours" and then --

MEMBER MOFFATT-BRUCE: And with this part.

MEMBER ADELMAN: -- this sentence over here, "Exclude" -- it's almost like including and excluding almost, in this sentence, are so similar. I'm not smart enough to -- it's almost like we're saying include them and exclude them.

I don't know if everybody sees what I mean.

MEMBER SMIRZ: Okay, E --

MEMBER MOFFATT-BRUCE: Yes.

MEMBER SMIRZ: It's include participants living at home --

MEMBER MOFFATT-BRUCE: And word, not.

MEMBER SMIRZ: -- as in the first site.

MEMBER MOFFATT-BRUCE: I think you're

missing a "not" in there. Right, so --

MEMBER SMIRZ: I'm sorry. Go ahead.

MEMBER MOFFATT-BRUCE: I think the questions are over here. "Were not identified less than 24 hours" --

MEMBER ADELMAN: Oh, that would --

MEMBER MOFFATT-BRUCE: I think -- is that --

MEMBER ADELMAN: That --

MEMBER MOFFATT-BRUCE: Right? Because you want to exclude, you want to only see them if they have it after 24 hours, right? Am I --

MEMBER ADELMAN: Okay.

MEMBER MOFFATT-BRUCE: Is that --

MEMBER ADELMAN: Yes, well that -- because otherwise you'll have the same, a very similar sentence for --

MEMBER MOFFATT-BRUCE: Inclusion and exclusion.

MEMBER ADELMAN: Yes, it just didn't. Okay.

MEMBER SMIRZ: Okay. So the first one is include participants where pressure ulcers were identified after 24 hours from returning from a congregate care setting.

And the exclusion is the same idea, so exclude them if they develop, in the home, less than 24 hours after they arrived back from a congregate care setting.

MEMBER ADELMAN: So it says less than up there but you just said after 24. But you said after, but it says less than. Maybe that's --

MEMBER MOFFATT-BRUCE: It's confusing.

MEMBER ADELMAN: For the inclusion you said something different than what it says.

MEMBER SMIRZ: Mark, do you want to do that? The -- so on the top portion, include participants living in home or assisted living facilities, include participants with pressure ulcers that developed and were -- "not" should be in there -- "not" is a typo -- not identified less than 24 hours.

Right. You can say it without double

negatives, yes.

CO-CHAIR SEPTIMUS: So what you're really trying to do is get an attribution --

MEMBER MOFFATT-BRUCE: Correct.

CO-CHAIR SEPTIMUS: -- so that if it develops within 24 hours, then you really can't attribute it to the PACE program.

And it really goes back to what happened before they were entered into the PACE program, as I -- so we go through the same stuff with HAIs except we usually use 48 hours, but you're using 24.

MEMBER QUIGLEY: But it was my understanding when this was first presented that this was, what was excluded was anything that developed after 24 hours of admission into the emergency department or the nursing home or the hospital because then it would be acquired. It would be associated with that admission rather than in the PACE program.

MEMBER MOFFATT-BRUCE: Right.

MEMBER QUIGLEY: So this should have stayed less than 24 hours for inclusion rather than

modifying that.

MEMBER ADELMAN: Yes, there's other poor language in there. Like I think you mean, in that same sentence, it says, "admitted to the hospital" and I think you mean -- I think -- "discharged from the hospital" meaning like how are we going to know if something happens before or after 24 hours after they're admitted to the hospital for ten days if you're doing all your evaluations at home?

So you mean after, if they were in the hospital and then they came home. I think.

MEMBER SMIRZ: You're right. PACE programs make an effort to obtain the clinical records from the hospital for the patient, to include them. So if they were, if they developed a pressure ulcer --

MEMBER MOFFATT-BRUCE: In the hospital.

MEMBER SMIRZ: -- after 24 hours in the hospital, that would meet the current CMS criteria for hospital-acquired pressure ulcers.

And so this definition follows that. So if the pressure ulcer happened essentially in the first 24 hours, it was present on admission, basically. And if it developed after that, then it was hospital-acquired.

MEMBER MOFFATT-BRUCE: It's just confusing. It's confusing.

CO-CHAIR SEPTIMUS: Albert.

MEMBER MOFFATT-BRUCE: They're in the PACE in the hospital --

MEMBER WU: Thank you, Dr. Septimus.

MEMBER MOFFATT-BRUCE: -- then PACE --

MEMBER WU: So is this, was it the reliability of this entire procedure that was tested? Or was it simply the judgment of an ulcer as Grade 3 or more or presence of Grade 3 or more?

Was it this whole algorithm which includes the inclusion/exclusion criteria that was shown to be moderately reliable?

MEMBER SMIRZ: No, it was the score. Whether it was all pressure ulcers or pressure ulcers Stage 3 and 4, it was just the scores that

were tested, not the procedures or the inclusion or exclusion criteria. Those were assessed by the validity team.

MR. LYZENGA: So again, this is similar to the previous measure where you're looking at the measure score. Essentially you're looking at the measure score's ability to distinguish between facilities.

And you're not really getting down to the data elements of whether those data elements are valid or reliable in themselves, but only if the measure score itself is able to distinguish performance across facilities, if it's getting just noise or if it's actually getting a signal that's telling you something about performance across facilities.

That's -- I'm not the right one to explain this. I wish we had our methodologist here who could talk about it a little more. But that's what measure score reliability is telling you.

CO-CHAIR SEPTIMUS: No, but I think we're looking more signal-to-noise in this

particular, in the reliability issue. Yes.

MEMBER WEBB: Just a picky point but, so one of the exclusion criteria is that you have to be in the PACE program for one day. But you can't --

CO-CHAIR SEPTIMUS: Per quarter. It's per quarter.

MEMBER WEBB: Per quarter, right. But you can't -- so a patient who was in there for one day, you can't identify a new pressure ulcer in that one day anyway because it's only after 24 hours.

MEMBER SMIRZ: Similar -- this follows the previous hospital issue. If the pressure ulcer was acquired in the first 24 hours, then it was not PACE-acquired.

So the people that were in for, have to be in for more than one day to be included in the count.

MEMBER ADELMAN: Yes. In some places it seems you're testing the reliability for Stage 3 and 4 but, correct me if I'm wrong, the measure is all stages. And for me, like NDNQI is Stage 2

and above. For me, Stage 1 would be, you know, the one where you have the most false positives, the little red spot and people are calling whatever.

Why do we test for reliability with Stage 3 and 4 but the measure is all stages?

MEMBER SMIRZ: In Table 2 --

CO-CHAIR SEPTIMUS: Look at the table there, first of all, Jason. There's, the top line is All. The second line is 3 and 4. 3 and 4 clearly have better reliability scores than All.

But the All still falls into the moderate range. That's correct. It's not -- the measure's All.

MEMBER ADELMAN: And I would be particularly concerned that Stage 1 will make it much less valid because it's, you know, a little red spot and you're -- you know, who knows what that is?

CO-CHAIR SEPTIMUS: And probably a little bit of subjectivity associated with that as well, I think.

Okay, I think -- did you want to speak again or just? Okay and so -- so let's go ahead and

vote on reliability -- oh, I'm sorry, Steve.

MEMBER LAWLESS: Yes, one question I just asked you. On the exclusion, exclude participants who are not in their home setting for at least one day --

MEMBER SMIRZ: Yes?

MEMBER LAWLESS: -- does that mean --

MEMBER SMIRZ: It means they either entered the program that day --

MEMBER LAWLESS: Right.

MEMBER SMIRZ: -- or they were in a nursing home for the entire quarter.

MEMBER LAWLESS: For all 90 days?

MEMBER SMIRZ: Yes.

MEMBER LAWLESS: So they had to be in for all 90 days. But if they left to go to a hospitalization for a day, came back, they would be excluded?

MEMBER SMIRZ: No. They would -- so if they were there. If they were in their home for less than a day we could not determine if they had a pressure ulcer, if it was PACE-acquired or not.

So people that were there for just one day are excluded.

So, but it also is for people in their homes so that people that are out of their homes for the full quarter are excluded. But if they were hospitalized for a short period of time, they'd still be included.

MEMBER LAWLESS: But just -- it says exclude who were not in their home setting for at least one day. So if they were hospitalized for a day, for two days in a quarter, they would be excluded or not?

MEMBER SMIRZ: No, they'd be included because they were in their homes for then 88 days.

MEMBER LAWLESS: Okay.

CO-CHAIR SEPTIMUS: Okay, well let's vote on reliable --

MR. LYZENGA: Hold on. Hold on one second. Sorry. And we actually just pulled in our resident methodologist. She's going to try to help explain maybe a little bit better than I do what we mean by a signal-to-noise test of reliability of

the measure score which is important because we've seen a couple already.

And we're going to see some more during the course of the next day and a half.

CO-CHAIR SEPTIMUS: I want to defend Andrew though. I think he did a great job. And he's doing it on much less sleep than most of us. So don't put yourself down, Andrew.

MS. JOHNSON: Okay. So I know Andrew did a great job. He always does. So I'll probably end up saying pretty much the same thing that Andrew did.

The idea of testing the reliability of the measure score is we want to be able to know if we can actually distinguish providers or not, right? So that's why we're doing it.

So signal-to-noise, the idea is you want to know how much of the variation that you're seeing in scores has to do with the differences between providers compared to the differences because of patients or because of measurement error.

So a signal-to-noise analysis looks at

that ratio, sorry, that ratio of variance that is between -- sorry -- variance between providers to the variance total, okay?

And that total, again, is the between and the within. And that takes into account measurement error. So it's just -- sure, sorry. Yes, kind of stuck here. Here we go. Okay.

So again, the signal-to-noise ratio is really a ratio of the variance that you see between providers to the variance overall. And that overall variance, again, is the between plus the within variance.

So I don't know if that helped you or not.

MEMBER MOFFATT-BRUCE: But using those numbers right there, help us interpret what those numbers.

MS. JOHNSON: Okay.

MEMBER QUIGLEY: Can you see it?

MS. JOHNSON: Yes, I can see it. It just takes me a minute. So when you do a signal-to-noise analysis, and I'm assuming this is

the Adams Method Beta Binomial. Okay.

What you get is an estimate for every provider, okay? So you get a number of reliability for each provider. And what they've done here is they've -- looks like they've shown you the mean and standard deviation and then the range and the medians. So those are the summary statistics.

So you can see that, on average, for that first one, the average reliability across all of the providers that they included in their testing, is 0.73.

So you would interpret that as -- Andrea, help me out. I feel like I'm on the hot seat here. I would think of that as 73 percent of the variation is due to variation between --

MEMBER MOFFATT-BRUCE: Providers?

MS. JOHNSON: Between providers. Correct.

(Off microphone comments.)

MR. LYZENGA: The sites, facilities.

MEMBER MOFFATT-BRUCE: Yes, but I just want to be clear, PACE sites? PACE sites.

MS. JOHNSON: Yes, sites. Yes. Usually, I think you have, we often use the term provider as kind of a generic term. So sometimes a provider is a hospital or sometimes it's a clinician or sounds like it's sites in this case.

MR. LYZENGA: And so, and I'm glad Karen's here. Maybe I can talk a little bit without -- and you can correct me if I'm wrong.

Again, as Karen mentioned, when you're doing this kind of analysis you get a reliability score for each of the facilities that you're analyzing. If you, and if you have a low reliability for each facility, for a given facility, that means you've got a lot of variation within that facility.

You're getting a lot of noise there. They're not getting a very consistent performance at the facility level. So you do a ratio of that to the overall variability, well the variation, both that and the variation across facilities.

What you want to have is a high ratio. You want to have a small amount of variation within

the facility --

MS. JOHNSON: And large across it.

MR. LYZENGA: -- and large across. So that you can --

MS. JOHNSON: Or you want relatively --

MR. LYZENGA: Relatively.

MS. JOHNSON: -- low variation within. See, if you have a lot of variation between your patients, that's okay as long as there's enough overall variation between to kind of --

MR. LYZENGA: Right.

MS. JOHNSON: -- overcome that noise, if you will.

MR. LYZENGA: So, and you're not really getting directly at sort of the reliability of the data elements, whether these things are beings things are being collected accurately.

You're sort of getting at it indirectly. Because if you don't have -- if you're not collecting the data in a reliable way, you're likely to have a lot of noise within the institution. Is that right?

So if you're getting high reliability scores at the measure score level, you're likely to be collecting the measure, the data in a reliable way. Although, again, this is not speaking about, to that directly.

It's more about the ability of the measure score to discriminate among measured entities and to distinguish good performance from bad performance.

DR. PINES: And just to make it a little simpler, so a good way to think about what sort of what's a good reliability number, usually think about 0.7 or higher is good for reliability. So both of these do meet that threshold, although the All ulcers rate is close to that threshold.

CO-CHAIR SEPTIMUS: Okay. If there's -- did you want to say something again, Kim?

MEMBER WEBB: So Steve and I were discussing this a little bit more. But I'm wondering if, in the denominator statement, it should include this for at least one day rather than an exclusion, right.

So I can tell you, I'm working now with our quality office. And these kinds of things confuse them so much that I think the report out would be not consistent among facilities and institutions.

And so I can tell you, if we're sitting in this room having trouble making sure that we understand it, I can tell you that quality officers around the nation are going to have just as much trouble.

And so we were talking about it. I think possibly using it in a denominator statement, the number of patients on a PACE organization census for at least X number of days, which I personally is two, not one during the quarter, could actually effectively get rid of the exclude persons who are not on the pay census for at least one day during the quarter.

Because I think that that, at least one day during the quarter, is confusing, is going to confuse the quality offices. And the reason I say two days is because you can't acquire a new pressure

ulcer in one day.

So having them only on the census for one day is a moot point, really, for this measure as far as I'm concerned. But that would just be a suggestion I would make because I think the reporting out of this measure is not going to be valid.

CO-CHAIR SEPTIMUS: So let me ask our experts. There seems to be some confusion over the specifications here and how it reads now. And we are voting on what is presented to us, correct? Not what --

DR. PINES: Yes, so -- yes.

CO-CHAIR SEPTIMUS: -- we wish to see but what's actually been in the actual measure itself. I just want to make sure that was correct.

DR. PINES: Correct. So what we're voting on here is that the specifications that were submitted, plus the reliability testing, if this does, you know, not go through, one option we could do would be to go back to the developer and ask them to revise those specifications for a future call

that we would put on again.

CO-CHAIR SEPTIMUS: Okay, so with that I --

MS. QUINNONEZ: We only have one non-consensus vote, correct?

CO-CHAIR SEPTIMUS: That's correct. We're going to probably look at this one too. So let's do it.

MS. QUINNONEZ: Voting is now open for the reliability of Measure 3000. Option Number 2 moderate. Option Number 3 low, and Option Number 4, insufficient. Option 1 high, Option 2 moderate 3 low and Option 4 insufficient.

All votes are in, and this vote is now closed. The reliability of Measure 3000 --

MEMBER ADELMAN: I just, for further account, there are three, who I think with this measurement, that we've been told are inaccurate or mistake.

And so like why are we voting on something that's just not even with the intent? Why don't we just put it aside, let them fix and

then we'll vote on it?

Because I don't even know what we're voting on. There's this exclusion that's the same as the inclusion. There's something about on admission was, I think it was supposed to be on discharge.

And then that last point, like if you weren't in your home for one day in a quarter then we throw you out. And nobody's -- I mean, unless I misunderstood that, everybody's going to be out of their home for one day in a quarter. If we exclude them you would exclude everybody.

DR. PINES: Well so -- so yes, so at this point, since it clearly didn't pass through this stage then this would go back to the developer to revise the specifications for a future call for a re-vote, should we want to do that.

MEMBER QUIGLEY: And if I may add, too, Pat Quigley, that in relationship to what Chris has said before, that this is something that CMS could be reporting already without NQF.

The pressure ulcer development is in the

PACE program, Type II error. They have Type I errors and Type II errors that they have to investigate.

A Type II error is defined as any pressure ulcer that develops in a PACE program. They have to do a 48-hour reassessment after that pressure ulcer has developed.

So that's where the staging should be done. You know, it's already existing in their program that's been around since 1997. So here we are, you know. So I just want to say that.

CO-CHAIR SEPTIMUS: Okay, so is there a recommendation that we refer this back to the developer and stop here? Is that what I'm hearing?

MEMBER MOFFATT-BRUCE: Yes, it's been --

CO-CHAIR SEPTIMUS: Is that? I'm not making -- I'm not --

MEMBER APPLEGATE: Yes, I would move that we go back to the developers with the recommendations offered today and that we record this in our notes or minutes.

CO-CHAIR SEPTIMUS: Is there a second to that?

MEMBER MOFFATT-BRUCE: Second.

CO-CHAIR SEPTIMUS: Okay, now discussion, Jason. So just --

MR. LYZENGA: We didn't talk to validity. And I just want to make a point today. I go back and work on the --

MEMBER MOFFATT-BRUCE: Mic.

MEMBER ADELMAN: We didn't talk about validity, but I wanted to make a point because I think I'm going to go back and work on it perhaps.

And it was hard to follow, but I think, from validity, all that really was done was experts said, you know, they thought it was valid. Like nobody looked at, for example, within NDNQI, we don't rely on staff nurses in the hospital.

I think if you really strictly follow the criteria, then once a quarter wound care nurses go around because they can measure the ulcers much more effectively.

So I'm just not sure if expert opinion

-- a bunch of experts saying we think this is valid, that's almost like our job. I would be much happier if somebody actually -- experts compared their read of pressure ulcers compared to what the nurses are doing and come up with the same outcomes. That would be valid.

MR. LYZENGA: So, Jason, I think that's, I think, data element reliability, or I mean validity rather. We do accept face validity, something like a technical expert panel giving their -- doing a systematic assessment of whether they think the measure score is valid.

It does give us a ceiling of moderate. We would, if we were voting on this, we would only have moderate as the ceiling for that vote.

I might also, if I could, I'm a little unclear on what our next steps are here for if we're going to sort of table this. And I might suggest, actually, that since we have consensus not reached here on these previous criteria that we actually just move forward and vote on the remaining criteria.

Because we will revisit these, as with the consensus not reached status, eventually anyway. Otherwise, I'm not sure what we would, if we would just address this on a post-meeting call and do a complete re-vote. Is that the idea?

CO-CHAIR SEPTIMUS: Okay, just to follow Robert's Rules so we'll have order this week, we have a motion that's been seconded, and we're now in the discussion period.

MR. LYZENGA: All right.

CO-CHAIR SEPTIMUS: Doesn't mean that we have to accept the motion, but I just wanted to make sure where we are with the discussion. All right?

MR. LYZENGA: Fair enough.

CO-CHAIR THRAEN: So I think one of the struggles we have is that when the reliability and the validity isn't working for us, it's hard to move forward into those other areas.

MEMBER MOFFATT-BRUCE: Correct.

CO-CHAIR THRAEN: And so even though I understand the push to move forward, it's like we're

all sort of sitting here going, well, I can't really make that judgment until this other piece is finished and completed and I feel like this is the right foundation to making these other decisions. So I think that's the struggle.

MS. MUNTHALI: And I just wanted to clarify, it did fail on reliability. I don't think there was consensus not reached. Am correct, Desi? So --

MR. LYZENGA: It did fail.

MS. MUNTHALI: -- I think the question is, it sounds like there is, the Committee feels like they'd like to see this measure go forward, and there may be some minor things the developer can do in the process by the post-comment call.

And I want to just get confirmation from the developers about their ability to be able to do some of the minor revisions that were outlined by the Committee.

MEMBER SMIRZ: Certainly. We can edit this quickly into something that's more easy to read, which I think -- I think that's contributed

to a lot of this discussion, that it isn't easy to read.

MS. MUNTHALI: Okay.

MEMBER SMIRZ: But it is complex. And then outline some, more clearly, some of the structure of the PACE program which the variance in that setting, from other kinds of settings, is also a complicating factor.

MS. MUNTHALI: So this is part of our process in which the developer would come to us for reconsideration because this measure essentially was not recommended by the Committee because a must-past criterion wasn't reached.

And so that would have to be done by the post-comment call. And you would discuss it then, and then you'll go on and continue vote, including an overall vote on this measure.

MEMBER QUIGLEY: Excuse me. I'd just like to say that I respectfully disagree on that. I don't think these, this discussion's points have been minor. I think that they are significant in terms of their scientific merit and integrity of the

measure coming forward from a program that's been in place for some time.

And I would like to have members look back at the pressure ulcer rates that came from the American Nurses Association, NDNQI, and the amount scientific rigor that was required for them to come forward.

So I don't think that this is minor. And I would just like to suggest than. Thank you.

CO-CHAIR SEPTIMUS: Yes. Okay, yes. I want to wrap this up. And someone can call the question for the motion. But go ahead.

MEMBER SMIRZ: As the developer of the NDNQI measure, I can say that the inter-rater reliability that you were talking about is easier to do in a congregate care setting than in, across people's homes.

And so it may be a feasibility of doing that kind of study. Certainly there is variation in the Wound and Ostomy Nurses Association on 24 hours, 48 hours for the development of an ulcer in another setting that shows up in a second setting.

And so if you have some of this -- if you have serious issues other than those, on your call if you would let us know what they are, we can speak to those.

CO-CHAIR SEPTIMUS: And I can't speak for everyone on the committee, but I think there's been enough discomfort with the measure as it is currently structured.

Plus, I think we haven't really discussed -- well, we did indirectly about whether or not, you know, Grade 1 is very subjective and should include all levels, or should it be just 3 or 4, which is what most of the other measures -- so I think there's more.

It's not -- I think we're uncomfortable with the measure as it is now. I think -- and I'm not sure. I can't speak for everybody, but I'm not sure that we feel as enthusiastic about this measure moving forward as we talked about a measure earlier today.

But I don't want to speak for the Committee. Albert and then Jason.

MS. QUINNONEZ: Lisa also.

MEMBER WU: So I just had a couple of --

CO-CHAIR SEPTIMUS: And Lisa. Sorry.

MEMBER WU: -- a couple of things that would make me feel sort of more enthusiastic about supporting this measure. And I do think it seems like it's terrifically important, that no one disagrees about that.

The first thing is that I think that I was a little unsettled about not having some data, some better data about the actual incidence, prevalence but perhaps even incidence, of what's happening in the program. And so I think if some, at least some data could be provided here, that would be great.

The second thing is is that while it's important to be able to judge an ulcer from not an ulcer and a higher grade ulcer from a lesser grade ulcer, it's important to figure out who's going to be doing it. And not everyone is equally good at doing it.

We know that physicians are probably

less good at it, from Susan's report. So I think that we would like to know in the setting that this is going to occur, can this be done reliably.

I think that we were unsure about applying the inclusion and exclusion criteria. And who you include in your denominator and numerator has a lot to do with whether or not you can do it reliably.

So I'd like some evidence that the procedure can be applied so that we get useful information.

And I think if you could get any kind of data on validity other than the content validity, face validity of those items, again, I would be reassured that we are looking at something that I believed was important.

MEMBER ADELMAN: I just wanted to reiterate that I think, what Albert just said, that I think pressure ulcers are important and they're -- I think this measure can be good.

They cause morbidity, and they're preventable. But in response to what Andrew said,

I understand that we sometimes use face validity. But I, for this measure as it is, I would push back that, like especially with Stage 1, there's just so much subjectivity.

And, I'm sorry, I forgot your name. But you had said that you actually felt the NDNQI measure that, as I recall, that starts at Stage 2. And there was some validity, reliability here done starting at Stage 3.

And I would -- if you really want to include all stages then I would, from my perspective, strengthen the validity testing or perhaps do what you did with NDNQI and start at Stage 2, because it's, I think, easier to identify than Stage 1.

You could even see, in your own reliability testing, that it got better as you had the more significant ulcers. Anyway, thank you.

CO-CHAIR SEPTIMUS: Lisa, did you put yours down? Lisa? Lisa? Did you -- okay. All right, so there is a motion on the floor. Hearing no other discussion, we will vote on the motion. Do

you want to re-state your motion?

MEMBER APPLEGATE: The motion is to table this measure until the developers revise it and bring it back to the committee with the recommendations that we've made or addressing them.

CO-CHAIR SEPTIMUS: Okay, I don't think we need to use our clickers for this. So all those in favor, say aye.

GROUP: Aye.

CO-CHAIR SEPTIMUS: Opposed? Oh, it's unanimous. So we thank you. You have another measure in just a second. But did you want to say one more thing about this measure?

MR. STEWART: Yes, just in closing, thanks very much for the feedback and recommendations. With CMS as the steward, everything we do is by taking direction from CMS. So we will take this back. Thank you.

CO-CHAIR SEPTIMUS: Tell them we love them. Okay, so the next measure is 3001, PACE Participant Fall Rates. I think you can quickly go through this --

MEMBER SMIRZ: Certainly.

CO-CHAIR SEPTIMUS: -- since we've already had this --

MEMBER SMIRZ: Okay, I'll take it section by section this time. The PACE fall rate is defined as the number of falls divided by the number of participant days.

So it's falls over exposure to falls. And so it's a ratio, not a percentage. The fall is defined as an -- looking for the definition -- an unanticipated descent to the floor or other surface where you would not expect to find a person and that -- and that a sudden, unanticipated descent in which the participant comes to rest of the floor or some other surface, person or object.

Inclusion criteria falls occurred in the patient/participant's home; if their home is an assisted living facility, in that assisted living facility if that's their usual place of residence; in the PACE center; or in the care of a PACE transportation operator.

So that the fall occurs in a setting

where PACE itself is responsible for the care. Participants who are assisted to the floor by a caregiver, are to be included.

Exclusion criteria, participants who fall, or let's say then sink, back to a bed, chair, car seat, walker, seat or toilet are excluded. So tried to get up, fell back down doesn't count.

Exclude falls in the participant home by other people, and exclude participants who are not in their home location or in the care of PACE in the settings I just mentioned.

So that's the definition of the measure.

CO-CHAIR SEPTIMUS: So this time -- first of all, I want to thank Susan for pinch-hitting on the previous, on very short notice. I think she did a great job. And, obviously, it was a very difficult measure to discuss. So, Susan, thank you for doing that.

This is one that I think Susan was prepared to discuss. So we'll turn it over to you.

MEMBER MOFFATT-BRUCE: True, I was prepared but I think some of the previous issues

come up in this one again.

CO-CHAIR SEPTIMUS: That's fine. So take us through the --

MEMBER MOFFATT-BRUCE: Absolutely. So the evidence is the first category. So with this, these are falls. These are all falls. I think that they have presented that, in other instances and other systems, that falls are a big issue and that we can impact them.

So in-patient, ambulatory settings. Not in the PACE. So they are abstracting, obviously, but there are issues that we can do and address to reduce the incidence of all falls, not falls with injury.

I'll open that for comment, for evidence.

CO-CHAIR SEPTIMUS: Comments about the evidence. Yes, Pat?

MEMBER QUIGLEY: Thank you. And thank you for bringing this measure forward. And I, while I reviewed the evidence to this, and I know we look at structure and process for fall rates, and

in recognizing how important fall rates are in this patient population, again, frail and in the home and in assisted living and nursing homes, the body of evidence that was presented to us is essentially acute care.

And I'd like to suggest that there is a whole body of evidence related to fall prevention in the home and assisted living that was omitted from this that looks at the importance of doing a multi-factorial assessment of not just the home environment but also the person.

But it also is related to the structure of care and the process of care. And the structure of care and process of care in the home setting is different than in the hospital setting, if you will.

So those bodies of evidence were missing from this review. And I submitted multiple comments of this in my notes in support of this program, of this measures as an outcome measure.

And predominantly, I'd like to say that one of the most significant components of preventing falls in the home setting and in home

care is the presence of an occupational therapist.

And that comes back to the importance of assuring fidelity of the program on the interdisciplinary team. And as I had mentioned, there was an analysis that was done in comparing the PACE program with home care programs.

And one of the limitations was ensuring the integrity of the interdisciplinary team. And there's no evidence to support how many of the sites that were included in this study, indeed, had occupational therapists.

If you will, as one example, if they actually have the structure to be able to implement this program in a home and assisted living program.

CO-CHAIR SEPTIMUS: And I should have mentioned at the beginning, this is an outcome --

MEMBER QUIGLEY: Yes.

CO-CHAIR SEPTIMUS: -- measure, and it's also a new measure. And I'm sorry I didn't make that statement first. Any other comments about the evidence?

Hearing none, I guess we'll vote on the

evidence. You want to speak? Of course.

MEMBER SMIRZ: I just want to mention that occupational therapists are mandatory members of the interprofessional team.

CO-CHAIR SEPTIMUS: Thank you. Okay, let's vote then on the evidence.

MS. QUINNONEZ: We are now voting on Measure 3001, PACE participant fall rate. And we are voting for the evidence. Option Number 1, yes. Option Number 2, no.

(Pause.)

MS. QUINNONEZ: Okay, voting is now closed. For the evidence of Measure 3001, 84 percent voted yes; 16 percent voted no.

CO-CHAIR SEPTIMUS: Performance gap. Susan.

MEMBER MOFFATT-BRUCE: Thank you. So I think this may be similar to previous, so the performance gap here, what we have demonstrated is again a sample of 50 sites from the 114 potential PACE sites whereby 34 submitted the data.

And they found a mean fall rate of 4.27

per 1000 participant days This, when you compare abstracted to hospital-based, it is high. And therefore the analogy -- the extrapolation is that there is a presumed performance gap.

I don't see evidence that there is a demonstrated or a calculated performance gap as it currently stands. I'd ask the developers if I'm missing something.

MEMBER SMIRZ: In the -- I don't know what page it is, but in the evidence of performance gap, there is a sort of embedded sort of table with the mean -- a standard deviation, mean, minimum and maximum for falls per 1000 participant days with a minimum of 1.88 and a maximum of 8.59.

So there is a substantial range. The mean is 4.27, and the standard deviation -- and the mean, the median is 4.4. So it's relatively normally distributed with a standard deviation of 1.53 which is reasonable.

CO-CHAIR SEPTIMUS: Any other comments on the gap? Yes, Yanling.

MEMBER YU: Thank you. When this

performance gap is compared with hospital settings, I don't know how to interpret it, because in hospital you have much better support system with the interdisciplinary teams to care for the patient.

So if we have found a lower score in comparing with the hospital, is this because your system isn't set up this way to really support, to prevent the fall?

Or is it due to the performance of this type of a PACE -- PACE, right? -- PACE setting, that enable you to compare from one PACE setting to the next, so there's intercomparability?

So they sound like, to me, like an apple and an orange when you compare them, to do this gap evaluation.

MEMBER SMIRZ: I think there -- it is apples to oranges with hospitals and PACE sites, although the measure definition is comparable. Because people in hospitals primarily are in bed more than people in their homes.

And so it might be expected that they

were, for the moment at least, more subject to falling than all PACE participants combined.

Second, the care processes are different. You don't have medical professionals in the home --

MEMBER YU: All --

MEMBER SMIRZ: -- 24 hours a day.

MEMBER YU: Right.

MEMBER SMIRZ: You do have occupational therapists that do go into every participant's home and eliminate fall risks such as throw rugs or poor lighting. They install grip -- hand-grip bars and other kinds of assistive equipment.

They check the participants for needed glasses as well which can contribute to falls. So those kinds of things are done which are different things that happen in the hospital.

MEMBER YU: So --

CO-CHAIR SEPTIMUS: You have another comment, Yanling?

MEMBER YU: So does that mean we should not compare them, in your opinion? To use this as

a comparison?

MEMBER SMIRZ: No, I think we -- I don't think so. Well, I mean, it's the same definition of measure. The care setting and the population differ. So you could say that fall rates are higher in PACE sites or in hospitals or in long-term care because the measure is the same, even though the setting is different.

They're probably not really comparable though. And so I think the important note is that there is a range of fall rates with PACE organizations or PACE sites.

MEMBER YU: Thank you.

CO-CHAIR SEPTIMUS: Okay, seeing no comments, we will vote then on performance gaps.

MS. QUINNONEZ: Voting is now open for performance gap of Measure 3001. Option 1 is high. Option 2 is moderate. Option 3 is low. And Option 4 is insufficient.

All right, all votes are in, and voting is now closed. For performance gap on Measure 3001, 11 percent voted high, 79 percent moderate,

5 percent low, and 5 percent insufficient.

CO-CHAIR SEPTIMUS: Okay. You feeling better now, developers?

MEMBER SMIRZ: Yes. A little better.

CO-CHAIR SEPTIMUS: Little better, huh? Now we're going to talk about reliability.

MEMBER MOFFATT-BRUCE: So relative to reliability, the inclusion criteria are for all calls including assisted falls, which I think is an important call-out to this.

Those that are excluded are those that fall back into bed, into a chair, car seat, walker or toilet or if they're not in their home location.

I have two, actually two or three questions on this aspect before I open it up. Firstly, it says that the excluded participants are when they're -- it's excluded when they're not in their home.

I presume that, though it must include when they're in transit to their clinics and such like that? Because I would think that that would be an opportunity to prevent falls for these

patients -- or participants.

MEMBER SMIRZ: Yes, that is included in the --

MEMBER MOFFATT-BRUCE: Okay.

MEMBER SMIRZ: -- as well as in the PACE center itself where they go for day services.

MEMBER MOFFATT-BRUCE: Okay, so that is included in -- wonderful.

And then the second question, I may -- I apologize if I'm asking this prematurely -- the source of the data. So is this type of fall, does it have to be agreed upon by the entire team that, indeed, it was a fall?

Is it anybody in the care team that can call it a fall? Who is, in fact, including it in the numerator?

MEMBER SMIRZ: Any person, any of the interdisciplinary team and the caretaker and the participant can report a fall. As long as it's documented in the clinical record, then we would count it.

The same is true of -- sort of true-- of

incident reports in hospitals where anyone, any staff member, can submit a fall which is then followed up by the risk care management team.

MEMBER MOFFATT-BRUCE: Okay. Those were my questions around the exclusion/inclusion criteria specificity.

CO-CHAIR SEPTIMUS: Okay, we'll put up the signal-to-noise data just in case, since I know we're getting really knowledgeable about this. But similar graph to the previous one. So --

MEMBER MOFFATT-BRUCE: Right, and so in this instance it's quite high. It's 0.83 but it has a large range, so that puts us into the moderate.

CO-CHAIR SEPTIMUS: Pat?

MEMBER QUIGLEY: Thank you. I would like to just say that I have some issues around the reliability or the inclusion criteria and the exclusion criteria for this measure.

In that all of the movement in relationship to fall rate is to not aggregate the fall rate. It is to go into precision about the type of fall.

We had this discussion when we were embracing the fall rate for hospitals, that we should look at accidental falls, anticipated physiological falls, unanticipated physiological falls. There's different types of falls.

But nonetheless, if this is the intent of CMS, when they have had a falls quality indicator in their Type II measures, it is a fall resulting in death or injury that requires hospitalization for five days.

We don't have any of that data that's been brought to us because those also have to have an assessment by the team within 48 hours.

To exclude a fall that because someone falls back onto a toilet is, I think, a mistake. Falls that result with -- associate with toileting result in severe injury, can be very grave.

Falls that are associated with someone going back into -- falling back into a chair or into a bed can also be grave. They can fall off of that. So there's no justification in the evidence to support the exclusion criteria. And again, this is

a very frail adult population, 55 and older.

In addition to the reliability and getting the signal-to-noise, the way that this was done in terms of aggregated data analysis is to group this patient population into only two groups -- those between the age of 85 to -- excuse me, 55 to 85 -- or 89, and those 90 and older.

And even the analysis of all falls data. when you look at one of the hallmark studies, original studies in 2015, published in January, we're using in the National Healthcare Statistics data to look at the aging population 85 and older in hospitals.

They looked at age groups by 10-year intervals. So I think that there's lots of issues surrounding the reliability of this variable and this measure. And I am totally in favor of fall rates, but with precision and with evidence to support it.

So as it's been presented today, irrespective of not having evidence that's associated with home care and assisted living, as

it's presented, in this variable, I think that there's a lot of limitations to it. Thank you.

CO-CHAIR SEPTIMUS: Iona.

CO-CHAIR THRAEN: So I'm going to counter you, Pat. So what you're speaking about is the maturity of measures and systems. And the hospital systems have been at this much longer than home care, home health, assisted living environments.

And so I think that the notion of moving towards precision in the world that you're talking about has come about because of starting somewhere to count falls. And in the process of counting falls, a certain amount of maturity has occurred and the realization that we need to look deeper and look more precise.

This is sort of a first-time effort in the home environment. And the level of sophistication is just not going to be there as you would find in a hospital setting as a starting point.

So I would argue for a learning curve

approach where you start first, just simply by capturing falls in this environment that has not normally been reported on.

And then, over the course of time, at least from the patient safety perspective, what I have seen historically is -- and I'll just give you a quick example -- when we started out reporting sentinel events, the professionals in the room, in 2001, were only willing to report eight general categories of events.

And then by 2005, as they learned that this was not adequate, it didn't support really truly understanding what was going on and they were more comfortable with capturing data, they moved that bar up to 32 specific events.

And so that the industry develops over the course of time, I think we have to give the home environment the same opportunity to grow and develop. That's my argument.

MEMBER QUIGLEY: And I would like to respond in saying that there have been measures through OASIS. And OASIS has been around for quite

some time. And that's the home care setting.

CO-CHAIR THRAEN: It hasn't always been home care health.

MEMBER QUIGLEY: But it's still the home care setting. And I'd also like to say that in 2016, with a body of evidence and the body of knowledge surrounding falls in today's world, it's very different than in 1997 when they started the program.

So the expectation of what is brought forward to us as a patient safety measure to evaluate the integrity of -- to improve practice and systems for an outcome that has severe consequences I think has a higher of expectation.

CO-CHAIR THRAEN: I just think we have to be careful not to be hospital-centric.

MEMBER QUIGLEY: And that's why I'm not. That's why I spoke to the evidence, that the evidence should have been grounded in the home and the home setting, not the hospital.

CO-CHAIR THRAEN: But home -- PACE -- the PACE people do sometimes have home health. But

what they really have is personal health care. And they have a tech or an aide that comes in and helps them bathe and helps them get out of the bed and move forward. It's not really --

MEMBER QUIGLEY: Which is why there should not be all these exclusion criteria if they're going to have this rate.

CO-CHAIR THRAEN: That arguing gets to the exclusion criteria as much as I'm arguing against the precision argument that you're making. That's the piece that I'm talking about.

CO-CHAIR SEPTIMUS: No cat fights here, sorry.

MEMBER QUIGLEY: Okay.

CO-CHAIR SEPTIMUS: One -- no, no, no, no. One, two, three, four. Go.

MEMBER WANG: I just want to throw in a quick spiel for, in support of starting somewhere in terms of reporting. The PACE organizations are small. Their participants are small.

When we -- if we get to, too quickly to the granularity, we may not have as much information

to share. So I just want to support starting somewhere.

MEMBER APPLEGATE: Oh, I actually did have a question similar to what Pat brought up, because I'm not an expert in this. But I do have family members who have fallen in the home, most recently last week.

And it is often related around bathroom issues. And so I did want to understand, or help me understand, from the developers not from anybody else, answer to the question about why that issue was excluded from the metric.

Because I think it is a really important part of -- going to the bathroom in the middle of the night, already having meds onboard, being a little bit dizzy when they got up, falling. Thank you.

MEMBER SMIRZ: Yes, that's my future. So, but, no, I do have really quite valuable concrete evidence about why we didn't include them.

One is that they're very difficult to record, unless they have an injury. And so in the

NDNQI, for example, which is an approved and endorsed measure for -- by NQF, includes sink back events if they involve an injury, but not otherwise.

We did a reliability, validity study on falls as recently as three or four years ago in which we did 20 videos of fall situations and then asked everybody, like 500,000 people in hospitals that could report a fall, if it was a fall or not.

And the sink-back incidences into a chair, sink back into, fall over on the bed, sink back to the toilet, unless we described it as involving an injury, was like a 50/50 split on whether it was a fall.

So by including those, you reduce the reliability and the validity of the measure.

CO-CHAIR SEPTIMUS: Yes?

MEMBER WU: Quick comment. I mean, I think that this is an argument for precision in general. And I think that if we start out by being very clear about what it is that we are after, I think that people in the home can do as good a job as we can in the hospital. So I'm arguing for

precision.

MEMBER ADELMAN: I just, I have a question again, I'm sorry, about the, just the specifics around the inclusion/exclusion numerator/denominator in that this is falls per 1000 patient days.

And the 1000 patient days, well, it's among patients who are at risk. It says something like that. And, to me, I know a lot more about falls in the hospital where risk is very well defined. I believe that there's also risk assessment tools for ambulatory as well.

And so I'm confused by, like the denominator of who's -- who are we talking about and who are at risk. And also, how do we get to the patient days? Like it is just -- well, let me let you answer that.

MEMBER SMIRZ: In answer to your first question, virtually all PACE participants are at risk of falling.

MEMBER APPLEGATE: Correct. It's a part of being in PACE.

MEMBER SMIRZ: Right. Yes, so they all receive risk assessment appraisals and then activities to reduce the risk of falling in the home.

Secondly, I apologize. I failed in the introduction to mention how we get per 1000 participant days. What the instructions say is to count the number of people in or enrolled in the PACE program in their home location every day of the quarter and then add up those numbers.

So it's the case load by day added up for the course of the --

MEMBER ADELMAN: Were they in the home?

MEMBER SMIRZ: -- yes. Or home-like setting.

MEMBER ADELMAN: How, if somebody was in a hospital for a month or two months, how does -- how do you know -- just how do they know that?

MEMBER SMIRZ: They know that because they pay for the hospital stay.

MEMBER ADELMAN: I see. So if it's --

MEMBER SMIRZ: So PACE -- one of the

things I haven't mentioned is the financial model behind PACE, which is basically an HMO. So they're responsible for paying for all care.

And so they would know that they were in the hospital probably both from their clinical records as well as their billing records.

MEMBER ARDIZZONE: I'm sorry. I just have a quick comment. I wanted to support Tracy, that we have to start somewhere. I know this has been around since 1997, but my understanding is there's nothing publicly reported about PACE right now. Is that correct? I mean, that's --

MEMBER SMIRZ: That's correct.

MEMBER ARDIZZONE: -- we need to start somewhere. We need to go somewhere. There needs to be public information out there about this program and about performance.

CO-CHAIR SEPTIMUS: Seeing no other hands, I think we're ready to vote on reliability.

MS. QUINNONEZ: The voting is now open on the reliability of Measure 3001. Option Number 1 is high. Option Number 2, moderate. Option

Number 3, low. And Option Number 4, insufficient.

Okay, all votes are in, and voting is now closed. For the reliability of Measure 3001, zero percent voted high, 89 percent voted moderate, 5 percent voted low and 5 percent for insufficient. Pass.

CO-CHAIR SEPTIMUS: We're on a roll. Okay, so the next one, I believe, is going to be validity.

MEMBER MOFFATT-BRUCE: So for validity, the testing that was used here is an example of the face validity with a consensus panel. The measurement strategy is dictated or is demonstrated by the ICBI which, in this instance, is 0.92, which is high.

The exclusions, again, and I might get some clarification here, were not tested because they feel that the exclusions are very straightforward. And then the, just under the threats to validity, and I had a question for the developers, the only risk stratification is age and gender.

There are no other factors that go into this. And yet those are weekly correlated with the performance around this measure. So I guess I would like a little bit of an understanding about what was trialed and if there are other, if there are plans to further develop some of the risk stratification.

MEMBER SMIRZ: Yes, certainly we're committed to, once we start collecting data from all PACE sites, looking at some socio-demographic adjustments. I know that those are difficult.

But more importantly, I think that there may be physiological adjustments to be made. But they also depend on what we find out from reliability and validity studies after they've been actually implemented and fully used for collecting the data.

CO-CHAIR SEPTIMUS: Just to let you know, since this is only face validity, we're not voting on number 1, just to remind people. Iona?

CO-CHAIR THRAEN: This is only moderate and --

CO-CHAIR SEPTIMUS: That's correct. It's only moderate or lower. So any other comments on it because it's only face validity? Just to make sure you remember that. So, seeing no other comments, we will vote.

MS. QUINNONEZ: We are now voting on the face validity of Measure 3001. Option Number 1 is moderate. Option Number 2 is low. And Option Number 3 is insufficient.

CO-CHAIR THRAEN: Thank you for naming them.

MS. QUINNONEZ: Option 1 is moderate. Option 2 is low. And Option 3 is insufficient. Thought that would make it easier.

CO-CHAIR THRAEN: Thank you.

CO-CHAIR SEPTIMUS: I thought Albert wanted to vote for Number 1.

CO-CHAIR THRAEN: I'm sorry, could you --

MS. QUINNONEZ: Absolutely.

MEMBER WU: Yes, you just -- you removed --

CO-CHAIR SEPTIMUS: Tell us when you want us to vote again.

MS. QUINNONEZ: Okay, voting is now open for the validity of Measure 3001. Option 1, moderate. Option 2, low. And Option 3, insufficient.

CO-CHAIR SEPTIMUS: No Number 4, Albert.

MS. QUINNONEZ: All votes are in, and voting is now closed. For the validity of Measure 3001, 79 percent voted moderate, 21 percent voted low, and zero for insufficient. Pass.

CO-CHAIR SEPTIMUS: Feasibility.

MEMBER MOFFATT-BRUCE: So around feasibility --

MEMBER DANFORTH: Can I --

CO-CHAIR SEPTIMUS: Time out. Was there a comment?

MEMBER DANFORTH: Yes, I just wanted to make one comment. Depending on how this goes, and I want to make it now before I forget, I do think because of the way the measure's specified and has

been tested, if it gets through this process and gets put into use, it'd be really important during maintenance when you come back to talk about any additional reliability and validity testing you did.

I think because this measure can't be validated through claims because these are falls without injuries, there's a high risk of underreporting, specifically not documenting when patients say that they fall.

I know, obviously, we're not there yet. But I did just want to make sure I made that point now. So, again, when you come back, it won't be as --

CO-CHAIR SEPTIMUS: Good point. If you can tell us how to get around that, we'd all like to know. But that's a great point. It's a great point. Okay. Feasibility.

MEMBER MOFFATT-BRUCE: So going into feasibility, which I think actually, Missy brings up with some good points. So this is captured from a variety of different sources.

We know in the in-patient setting, it's difficult, and we have a lot of ways to capture data. I can imagine that there may be some challenges with this. It does speak to having studied the sites thus far and that there tends to be a little bit of a learning curve to capture some of this data.

But I'm wondering if the developers may comment on how we can make this easier for the sites to actually capture validity, valid data in a reliable fashion to make it feasible.

MEMBER SMIRZ: Well, first of all, we will, if it's implemented for public reporting, we will -- or any accountability purpose -- we will do training of the sites and provide sustained resources for updating new people as they come into the organization on how to collect the data.

We may make it mandatory that no matter where it's discovered that it be included in the clinical record so that it can be extracted from the healthcare record.

And as for undercount, we have talked about giving caregivers who are in the home every

day logs to record falls so that it's not just the self-report of the participant, who's afraid that potentially they may be put in a nursing home because of a fall, whether or not there was an injury.

Undercount is, of course, a serious issue --

MEMBER MOFFATT-BRUCE: Sure.

MEMBER SMIRZ: -- for home care.

MEMBER MOFFATT-BRUCE: The other question I have, just for my own edification, there's 114 PACE sites. What's the ratio of electronic versus paper documentation in these sites?

MEMBER SMIRZ: We don't know that. Actually, I think they're in the process of developing electronic health records. Many of them have them. Whether or not this data -- we asked in a post-data collection survey to ask how data collection went.

MEMBER MOFFATT-BRUCE: Sure.

MEMBER SMIRZ: We, most of them

hand-extracted their data from --

MEMBER MOFFATT-BRUCE: From various sources, right.

MEMBER SMIRZ: -- electronic records. So what they really need is a way to get it programmed in on an ongoing basis, which they would likely do if was mandatory reporting.

So, but there's still PACE sites that are still using --

MEMBER MOFFATT-BRUCE: Paper?

MEMBER SMIRZ: -- paper. Yes.

MEMBER MOFFATT-BRUCE: Okay. Thank you.

MEMBER COOK: Yes, on both of these last two measures, it was stated that there was 50 sites that were randomly chosen for the data collection going into the validity and reliability testing.

But yet, in both cases, you've had somewhere around 30 or just short of that. Does that give you any indication or did you get feedback from those facilities which did not actually supply data?

And do you think that will be a hindrance to the feasibility of this measure being implemented?

MEMBER SMIRZ: We examined the program age and program size and geographic location of the programs that did not participate along with those that did. And there were no significant differences so -- or very minor differences.

So we didn't have any evidence to support bias. The unmeasured thing, of course, is some programs may just find it easier to access the data. We didn't have information on that. So it's possible, but not that we found out.

MEMBER LAWLESS: Curious. In terms of performance management evaluations of the people who are in the home -- oh. In terms of performance management evaluations for people who were helping, are they uniform throughout PACE?

I mean, so everybody who's at home is judged in terms of their performance the same? And is this part of that performance, reporting it or falls at home?

MEMBER SMIRZ: I have no idea. I have no idea. Tamika, are you still on the phone? Possibly not. We'd have to look into that and get back to you.

MEMBER LAWLESS: Because if it's uniform, and it's not, their performance is not tied to whether they're either reporting or ER visits related to a fall, you can see where it --

MEMBER SMIRZ: Sure.

MEMBER LAWLESS: -- you would have a discrepancy.

MEMBER SMIRZ: Right. Good idea.

MEMBER ADELMAN: This is just a small point, but I just -- it just occurred to me that when I asked you before about the number of days, the denominator as in patient days, and you explained it to me.

But I don't actually think that that's in the measure. I just think you should add the language so that those who read the requirements and want to do it know the exact rule. Unless I missed it.

MEMBER SMIRZ: It was in the very back section of the 40 or 50 pages for the measure but not in the up front description. So, but yes, we will make sure that it's added. And I think the language is fairly clear.

MEMBER ADELMAN: I think that the true denominator is patient days. That's the denominator. And it's patient days when they're home, according to what you told me.

MEMBER SMIRZ: Right.

MEMBER ADELMAN: And that is the part that's not clear. It's --

MEMBER SMIRZ: Okay.

MEMBER ADELMAN: When you explain it, it's clear. But I think the denominator statement should reflect that.

MEMBER SMIRZ: Okay.

CO-CHAIR SEPTIMUS: Chris, you've had -- you still want to have a -- see if any -- is that still -- all right. So I don't see anymore hands so we will vote on feasibility.

MS. QUINNONEZ: Voting is now open for

the feasibility of Measure 3001. Option Number 1, high. Option Number 2, moderate. Option Number 3, low. And Option Number 4, insufficient.

We're looking for one more vote. All votes are in, and voting is now closed. For the feasibility of Measure 3001, 0 percent voted high, 74 percent voted moderate, 26 percent voted low, and 0 percent voted insufficient.

CO-CHAIR SEPTIMUS: Usability.

MEMBER MOFFATT-BRUCE: So this is a new measure. It's not currently publicly reported. There are plans, I understand, to put it into an accountability program.

And I think what we are hearing today is that there's a need and a desire to have this for these patients that are so vulnerable and where falls are such an important issue.

CO-CHAIR SEPTIMUS: So question to our NQF staff, because this is a new measure and there really isn't any experience with usability, how do we --

MEMBER MOFFATT-BRUCE: How do we vote?

CO-CHAIR SEPTIMUS: -- evaluate this section and then vote? Since there's really no data on usability, because it's not publicly reported and it's a brand new measure, how do we evaluate usability in this setting?

MR. LYZENGA: We have some planned use, it looks like, that they put in which is not unusual for newly developed measures. This one's a little bit more subjective than some of those other criteria.

It's basically just, if you feel that they have put in -- described a good enough plan to put this measure into use and to -- in quality improvement, public reporting or other accountability programs, then you should vote accordingly.

If you do not feel like they've given us enough information on how they plan to put this into use, then you can reflect that in your votes, I think.

DR. PINES: And also just to comment, so this is a potential usability, so especially for

future measures.

CO-CHAIR SEPTIMUS: That's, thank you. That's exactly what I wanted to hear. So let's start from the back and work up here on the left. This is the murderer's row here on the left.

(Laughter.)

MEMBER DEED: Okay, not yet but I really need to hear more from the developers about the potential for public reporting of this measure. I'm very uncomfortable with that checked box of not publicly reported. Thank you.

CO-CHAIR SEPTIMUS: Please comment.

MR. STEWART: So most of the comments we've made today are on behalf of the measure developers. CMS is the measure steward, and they'll be making all the decisions around implementation.

CO-CHAIR SEPTIMUS: Can we assume that obviously, since they've contracted with you to do this, as is the case in many of these CMS measures, that they in fact do want to have this publicly reported and potentially linked to payment, can we

assume that?

MR. STEWART: Yes, sir. The measure developer assumes that, especially the former. The reporting the latter linkage to payment is probably further down the road at CMS.

CO-CHAIR SEPTIMUS: So pay for reporting and then pay for performance has generally been their mantra. So let's keep going.

MEMBER WANG: So in terms of quality improvement, I would assume we wanted the ratio to go down, right? Lower rate the better.

But how do we just, since this is a new metric, how do we distinguish the difference between a higher rate of a fall rate because this is a new metric and more people are reporting versus it's a quality improvement? I mean, eventually it's going to go down.

MEMBER SMIRZ: I'm not sure I got your question.

CO-CHAIR SEPTIMUS: You want to repeat it?

MEMBER WANG: Can you hear me? Okay,

yes. So just a quality improvement, we expect the falls rate to go down eventually --

MEMBER SMIRZ: Right.

MEMBER WANG: -- with all of the proper implementation of best practices. But because this is a new metric, and I'm assuming that there will be -- to implement this, that people will be trained on how to report falls.

MEMBER SMIRZ: Correct.

MEMBER WANG: So theoretically, there might be an increase in the reporting of fall rates initially. And how do you distinguish that from a truly quality improvement effort to reduce?

MEMBER SMIRZ: You're right that when you start measuring something, things may go up. The, I think, key point is, actually, around use and usability, which is it's common practice, and I believe part of CMS's blueprint on measure development to have measures be collected and reported to CMS for some period of time without using them for public reporting or any other accountability purpose for a year or two.

And then if -- once they're stabilized, then the use for either of those purposes can occur.

CO-CHAIR SEPTIMUS: Not to bring up a prior discussion, but, in fact that's exactly what CMS is doing with the Sep 1 measures for sepsis. I don't want to get into that discussion though.

MEMBER QUIGLEY: Thank you. And in my remarks related to usability, it is still to always support the importance of fall prevention in any setting of care.

But I still would like to go on the record in public that to have such aggregated fall rate is not going to be a driver for improving patient safety in 2016, recognizing the amount of work that's gone on in other industries of health care for greater precision.

In long-term care there's much more relevant rates for this kind of a patient population. The percent of patients who fall that are in the care, the percent of repeat fallers, the percent of recurrent falls.

This fall rate, as presented, is so

aggregated that there's no way to be able to really link quality of care, of structure and processes, to the fall rate.

And in the remarks that have been made related to this rate, there has been said that this could identify a good program or a bad program. And quality measures are not to be able to criticize good or bad but to be able to profile risk and improve practice.

And this patient population is a falling population. They are a falling population, 55 and older, frail, older people. So I would just like to say in terms of usability, my expectation would be higher.

It is not to say that there should not be some starting point, but to be able to really drive quality and safety of care, I think that this is very limited. Thank you.

MEMBER WEBB: So I just had sort of a question about the practicality of this. I had never heard of PACE before we started doing this, and I can't imagine that there's more than one PACE

organization available in any given area, honestly.

MEMBER QUIGLEY: That's true.

MEMBER WEBB: At least in the areas that I have worked. And from a practical standpoint, if a PACE organization is failing and we're going to pay them less because they're failing, what are we going to be doing to patient care for those patients?

There's no alternative. So what would be the plan at that point? I mean, I'm just thinking about sort of the patient's viewpoint on this. You know, it's not like it's a capitalist adventure. You don't have anywhere else to go.

The alternative would be to put these patients all in nursing homes if we stop paying the PACE organization they can no longer support.

CO-CHAIR SEPTIMUS: Let's have the developers respond.

MR. STEWART: PACE organizations are by ZIP code. You're correct. But you could disenroll. You're choosing to participate in PACE as your care provider.

CO-CHAIR THRAEN: I was just going to comment that there is sort of a precursor to what's going on right now with the Medicare change towards value-base care versus volume-based care.

And so it's a managed care model that combines both Medicaid and Medicare streaming of funds. You have to qualify. You have to be eligible. You have to meet the criteria.

It was originally 55 and older. In the article I was just reading, in 2009, they were definitely looking at extending it to the younger disabled population, younger than 55. I don't know if that change has happened.

And with this movement towards managed care models or patient-centered medical homes or whatever it is you're going to call it, which is managed care, basically, on the part of Medicare and Medicaid, you're actually, I think, going to see more opportunity for folks to enroll in this kind of program.

It may not be, it may not in the long run be this PACE program. It might be patient-centered

medical home rebirth kind of thing. But they're definitely moving, Medicare is moving in that direction.

They've set very aggressive goals on how many people are going to be enrolled in these kinds of programs, meaning those of us sitting at the table are going to be in managed care organizations as we get older.

So I think that it's going to continue to move forward, and it's very specific to that combination Medicare/Medicaid population at this point in time.

MEMBER DEED: Yes, the reason why it is so important to have these measure publicly reported from the beginning is exactly what's been said here, which is often there is no choice. But if families know what's going on, they can help at their end to try to make it better. And the change can happen faster.

CO-CHAIR SEPTIMUS: Thank you. It's always great to have patient advocates to comment on these measures. All right, based on what our NQF

colleagues said about how we vote on usability, I think we're ready to vote.

MS. QUINNONEZ: Voting is now open for usability and use of Measure 3001. Option Number 1 is high. Option Number 2 is moderate. Option Number 3 is low. And Option Number 4, insufficient information.

Thank you. Okay, all votes are in, and voting is now closed. Voting for the usability and use of Measure 3001, zero percent voted for high, 82 percent voted moderate, 18 percent voted low, and 0 percent voted insufficient information. Pass.

CO-CHAIR SEPTIMUS: If I remember my order, we're ready to vote on whether or not this measure is suitable for endorsement by NQF. So this is an easy one, Albert. It's only 1 or 2.

MS. QUINNONEZ: Voting is now open for the overall suitability for endorsement of Measure 3001. Option number 1 is yes. Option number 2 is no.

All votes are in, and voting is now closed. For the overall -- excuse me, overall

suitability for endorsement, 94 percent voted yes, and 6 percent voted no.

CO-CHAIR SEPTIMUS: Okay. So, we're sort of at this tipping point. We have another measure which is very similar to this measure, by the same developers. The only difference in this one is it's with injury.

I think this will go a little bit faster. Or we could take a break now. It's up to you. Which one you want to do? We're going to -- definitely going to take a break. It's either or after the next one. Keep moving? All right. That's the -- we don't have to vote on that one. I think we have a consensus on that one.

Okay. So why don't you just maybe quickly tell us if there's any specific differences, and then Pat, I believe, is the discussant on 303. Again, this is a new measure and an outcome measure.

MEMBER SMIRZ: The only difference in this measure is that it includes only falls with injury in the numerator. We also have a

classification for types of injury, from none to death, which pretty much covers the range.

And all of the rest of the analysis for reliability and feasibility was then just conducted on falls with injury, as opposed to total falls.

CO-CHAIR SEPTIMUS: You know that we're, we've sort of, been more familiar with measures that talk about injuries, such as the HAC measure. So this would be perhaps more in line with what we discussed before.

So you already have questions before Pat even --

CO-CHAIR THRAEN: Yes. I want to ask --

CO-CHAIR SEPTIMUS: -- starts? Okay. Somebody's -- someone has a mic on.

CO-CHAIR THRAEN: Okay. There. So as you read this, through this, does this one carry the kind of precision that you were looking for before?

MEMBER QUIGLEY: Thank you for that opportunity to answer that question. The issue surrounding this is the lack of evidence to support

it, and the model that was selected to support the measure.

As an outcome measure, we are expected to look at the structure and the process for fall injury reduction. May I proceed?

CO-CHAIR SEPTIMUS: Go ahead.

MEMBER QUIGLEY: Okay. So that's the quick response.

CO-CHAIR SEPTIMUS: So Yanling, and then we'll -- oh, and then Lisa. I'm sorry. Then we'll go to the first question.

MEMBER YU: Thank you. Just a matter of helping me better understand this measure, if I understand correctly, the only major difference is this one is involved with harm, fall and, you know --

MEMBER QUIGLEY: Yes.

MEMBER YU: So is there any way, in your mind, that these two can be combined together with the last one, or this -- what is the rationale that you have to separate them, make different measures?

MEMBER SMIRZ: Right. No. I

understand completely what you're saying. And the -- previously for, let's say, NDNQI, there's been some discussion about whether to make these paired measures, not to combine them into a composite measure, but to make them paired.

And there was -- I'm not really sure -- I think of paired as something like a process measure with an outcome measure, as opposed to two outcome measures.

So this is an alternative. In general, I think that some more -- there's more of a consequence for a fall that involved harm, both for cost, for the discomfort and disability of the patient. And so some people prefer that measure to the other one. But you don't -- in general, what you do to prevent falls will also prevent the injury, if you prevent the fall.

So this is an auxiliary measure, sort of to give you sort of a level of harm. And there's also some thought that the falls with injury measure has higher reliability, which it does, slightly, because it's more likely to be reported.

CO-CHAIR SEPTIMUS: Well it may be -- if I -- I think, it may be like the near miss in drug. Is that a good analogy, then? So if you look at processes where people are falling without injury, that processes may -- if it's fixed, may prevent someone who has harm. I mean, isn't that the rationale? Okay, so Lisa?

MEMBER YU: Could I just add one --

CO-CHAIR SEPTIMUS: Yanling, I'm sorry. Are you finished?

MEMBER YU: Does that mean that, in term of public reporting, or especially about CMS accountability program, would that be a different set of incentive implementation for CMS, as far as you know?

MEMBER SMIRZ: I do not know their plans.

MEMBER YU: Okay. Okay, thank you. It just matter --

CO-CHAIR SEPTIMUS: Lisa?

MEMBER MCGIFFERT: I was going to ask about the definition of the falls, and then I found

the section where you, where it's discussed. But it is a pretty broad range.

And is that typical with falls, that when you're defining falls with injury, you might know this, but it -- that you don't really look at the type of injury? It's just like a minor one might require a dressing; a major one might cause death. There's a --

CO-CHAIR SEPTIMUS: I think we're going to get into that when we go through the, this --

MEMBER MCGIFFERT: Never mind.

CO-CHAIR SEPTIMUS: -- the measure. No, no. That's a great question, though. So it -- Jason, did you have something to say before we even discuss evidence? It's okay. Jason, is your mic on?

MEMBER ADELMAN: NQF endorses the common formats and has a definition of injury with falls. And then we also endorse the NDNQI that has a different definition of falls with harm.

And I was hoping that we wouldn't have a third definition, also endorsed by NQF. But the

third definition -- and actually, I was looking, and I didn't see any definition for harm, which I guess is like a third -- but I might have missed it.

But I think -- did I miss it? Because I didn't -- we'll see.

CO-CHAIR SEPTIMUS: Did --

MEMBER ADELMAN: But when we get to it, we can -- well which harm scale? The AHRQ one, or the NDNQI one, or the? Is that what -- is used here?

MEMBER SMIRZ: No, no. I don't think --

PARTICIPANT: NQF. It's the NQF severity rating scale.

CO-CHAIR SEPTIMUS: I think Pat is --

PARTICIPANT: It's the NQF.

CO-CHAIR SEPTIMUS: -- is going to discuss that. But so --

CO-CHAIR THRAEN: I just had a quick question. So when you said that this is the same as the one before, are the denominators the same? So before you have the -- you have the number of falls for the whole group, versus the number of

falls with injury, over what?

MEMBER SMIRZ: The whole group.

CO-CHAIR THRAEN: The same, same denominator?

CO-CHAIR SEPTIMUS: Anything else? I mean, as with the other discussion, we're going to discuss numerators, denominators and exclusions, so hopefully we'll get to that, but so with that -- are you finished, Yanling? Okay. All right.

So now Pat, you have the microphone.

MEMBER QUIGLEY: Thank you. Thank you, Mr. Chairman. Thank you for the opportunity to be able to present this indicator on behalf of CMS to this body.

And the first issue that we get to address is the structure and the process. And in addressing the structure and the process to be able to help reduce fall and fall-related injuries, I looked at the model that was presented in the submission to us. And the model that was presented is the same model that they have for being able to prevent falls.

So what is missing from the model is anything that is in place to assess risk for injury. And that would be, for example, osteoporosis or anticoagulation, or anyone who has already had an injury history that has occurred.

And it also does not include the processes to be able to reduce injury in the home setting or the assisted living care facility, because the interventions and the processes to prevent injury are separate and distinct from fall prevention. We just talked about fall prevention.

In that regard, since I discussed the structure and the process in terms of the model, that led me back to -- because I had to look at the model that was done to the analysis of that report, that it compared PACE with home care.

And they did have a model in that analysis by the Mathematica Policy Research, Incorporated in 2008, that had included in their structure, the integrity of the interdisciplinary team. And in this case, it would be to prevent injuries from falls, because colleagues, you cannot

prevent all falls. And this is a repeat falling population.

So having to go back and then look at the evidence, I did an analysis of the evidence that was presented to this body to support this measure. And on the screen, if we go back to the evidence, the Fong article is from Portugal, and has nothing to do with injury reduction.

I cannot find the Levonden (phonetic) article. The Rara (phonetic) article is from Sri Lanka, one district in Sri Lanka that does not address injury. And the other literature that is there is essentially the literature to support the fall injury literature in hospital-based.

Even though in the analysis of the population at risk, the outpatient population, there is a lot of evidence that's presented by the developers in relationship to the prevalence of injury in community-dwelling elderly by CDC. But the actual literature that we have here does not really support the measure in relationship to injury.

As an outcome measure, the -- we have the numerator that is presented, and a numerator is indeed falls with injury experienced by the --

CO-CHAIR SEPTIMUS: Pat, Pat, Pat --

MEMBER QUIGLEY: I'm sorry.

CO-CHAIR SEPTIMUS: Time out.

MEMBER QUIGLEY: Stay at injury?

CO-CHAIR SEPTIMUS: Let's go with evidence.

MEMBER QUIGLEY: Okay.

CO-CHAIR SEPTIMUS: We'll get to the others. I know we all want to jump ahead, but otherwise, we're going to get very confused. So just go by the evidence.

MEMBER QUIGLEY: But I did present a summary of literature and literature to support this measure that could have been more relevant.

CO-CHAIR SEPTIMUS: Okay. Any other discussion about the evidence?

MR. LYZENGA: And just to clarify, the -- I mean, the question again, on an outcome, is whether there is at least one process or -- for a

process, you want a systematic review and QQC, but for an outcome, all you need is a rationale, showing that there's something the provider can do to affect this outcome.

And technically, in this kind of situation, usually, you know, where you believe there is evidence that there is something a provider can do, but it's not provided by the developer, I mean, one option is to vote insufficient and then say insufficient with exception.

MEMBER QUIGLEY: Well that's the purpose of my discussion --

MR. LYZENGA: Okay.

MEMBER QUIGLEY: -- and presentation.

MR. LYZENGA: Or you could also, if you believe there's enough and want to present it to the committee, you could also -- again, all we need is a rationale. We don't need a lot of evidence for an outcome.

CO-CHAIR SEPTIMUS: Lisa?

MEMBER MCGIFFERT: So my understanding of this is that if we're looking for is there a

prevention for this, and there are certainly ways to prevent falls, but I think I -- I'm not sure that there are, is a way to prevent falls with injuries, because of -- because the injury is really dependent on the condition of the patient.

I could fall, and it couldn't hurt me at all. Someone else could fall, and it could make them disabled for life. There -- and that is -- so I have difficulty with this, as a -- I like the all falls measure better, because it's the patient's condition that determines the injury in -- not in all cases, I'm sure, but in many cases.

MEMBER QUIGLEY: Thank you for that comment, Lisa. And my response to you is that there are interventions that can be put in place to reduce injury. And that's -- a lot of that work has been done by the Department of Veterans Affairs for over 15 years, to go after injury reduction as the primary outcome, not falls.

So in the home setting, there are things that can be done to pad environments, eliminate sharp edges, to reduce impact of falls, and to

identify those who are vulnerable, those with osteoporosis risk factors, those who are anticoagulate, et cetera, which is why my remarks were: the model didn't fit the variable. The model that's still presented, it fits the fall rate.

CO-CHAIR SEPTIMUS: Kimberly?

MEMBER MCGIFFERT: Does that -- would that prevent -- if I had osteoporosis, that a padded -- like something padded would prevent that?

MEMBER QUIGLEY: So would we prevent your hip fracture, or hip protectors or helmets could prevent head injuries? Exactly.

MEMBER MCGIFFERT: Yes.

MEMBER QUIGLEY: Yes.

MEMBER MCGIFFERT: Okay.

MEMBER APPLEGATE: I think what Lisa's bringing up, though, is there's confounding error between the two measures. And I -- again, I would raise the question about whether both measures are necessary, or whether the second measure is going to overlap with the other measures that currently exist.

MEMBER ARDIZZONE: I just wanted to say, I think the two measures are both very important. We need to see how many times patients are falling in these PACE programs, and they need to be publicly reported, as well as we need to see, these are frail patients who are falling and injuring themselves. That is going to be bad for their morbidity and mortality.

So they're two -- I know they, it sounds like they're capturing the same data, but the effects on the patients are very different. I forgot my other point.

MEMBER ADELMAN: I just want to second that point. I feel the same way. I see how it can be confusing, but I also see -- and I agree. If you prevent all falls, then obviously you'll prevent falls with injury. But there are special things that we can do with people who have higher risk for injury.

And sometimes we don't have enough falls with injury to see the effect of an intervention, so overall falls helps to study how effects are.

And I could go on, how I see the value of both, but I agree with that point.

CO-CHAIR SEPTIMUS: Again, I mean, just to follow up on, or at least, I mean, you could look at -- not quite the same, but there are some parallels for HAI reduction. We know that there are some intrinsic patient factors, but we also know that the majority of them are preventable.

So the goal is to try to get the rate as close to zero as possible, but it may not be zero. So it may be we have a baseline rate. We put in intervention, and then we -- it's really more important to trend that over time, to see whether our interventions are, in fact, having the desired effect. So I don't know if that's with the spirit of this measure or not.

MEMBER SMIRZ: Yes. It -- yes. Trending and understanding both would lead, potentially, to different kinds of interventions, but also as I mentioned earlier, the patients are assessed -- are assumed at risk, and their environments are assessed for improvement.

CO-CHAIR SEPTIMUS: And there's probably some opportunities for research, in terms of the translational type of contextual issues that sometimes we don't often study.

MEMBER SMIRZ: Correct.

CO-CHAIR SEPTIMUS: So. Yes, Iona?

CO-CHAIR THRAEN: So when you -- using the overall fall rate, and the data that's intended to be collected from the PACE program, in that process of collecting that data, how is this measure -- what is this measure collecting that's different from that data?

MEMBER SMIRZ: It just adds to the field for -- well two fields, one, was there an injury; and two, what was the level of injury?

CO-CHAIR THRAEN: So in reality, when you pull the data, you're actually getting --

MEMBER SMIRZ: Both.

CO-CHAIR THRAEN: -- both at the same time and reporting it out as two separate notions? So there's not additional burden or anything like that.

CO-CHAIR SEPTIMUS: Kimberly, do you want to speak again? Yes.

MEMBER APPLEGATE: One more. One more thing. If patients are injured enough by these falls, how many of them stay in the PACE program, and how many are transferred to other care facilities?

MEMBER SMIRZ: Need to get the data to find out.

CO-CHAIR SEPTIMUS: Okay. Let's vote on the evidence.

MR. LYZENGA: And I should note that I was mistaken that there is an insufficient option. There is not for outcome measures, just yes or no. So again, is there a rationale supporting the relationship of this health outcome to at least one health care structure process, intervention or service, yes or no?

MS. QUINNONEZ: Voting is now open for evidence of Measure 3003. Option 1 is yes. Option 2 is no.

Okay. All votes are in, and voting is

now closed. Voting for the evidence of measure 3003, 95 percent voted yes; 5 percent voted no.

CO-CHAIR SEPTIMUS: Excellent. So the next thing would, of course, would be the gap in care. So we're looking, is there a gap in care that warrants a national performance measure. Pat?

MEMBER QUIGLEY: Thank you so much. They conducted the same analysis using the same process with the 50 sites and showing that there is an opportunity for improvement in identifying the fall rate, because in terms of performance and data that was presented -- we have data that's presented related to the falls, the number of people who fell as well as the number of patients who fell with injury.

So we have opportunities for improvement, and there is a performance gap to be able to reduce injurious falls. They have it presented as a total population, as well, and they have the statistics that are there to support that.

CO-CHAIR SEPTIMUS: Okay. I think Pat probably said it all. So, and seeing no hands --

MEMBER QUIGLEY: And that was short. You were surprised, weren't you?

CO-CHAIR SEPTIMUS: Hey Pat. So now we're going to go to vote on performance gap.

MS. QUINNONEZ: Voting is now open for performance gap for Measure 3003. Option number 1 is high; option number 2, moderate; option number 3, low; option number 4, insufficient.

All votes are in. Voting is now closed. For the performance gap of Measure 3003, 32 percent voted high, 63 percent voted moderate, 5 percent voted low, and 0 percent voted insufficient.

CO-CHAIR SEPTIMUS: Okay. Now I think we're up to -- where are we, reliability? Okay. So have they explained their rationale, and can this measure be consistently implemented reliably? So Pat.

MEMBER QUIGLEY: Thank you. The reliability is considered to be high. The reliability is the same model, the signal-to-noise model, but also in identifying the level of severity because they are using the severity rating scale for

injurious falls from NQF as it's presented in the report.

CO-CHAIR SEPTIMUS: You know, this actually has the highest signal-to-noise of the things that we've said. This is, if I read it right, is 0.88. So we're going to put that up on the screen in case -- now that we've all become experts on signal-to-noise. Any --

MEMBER ARDIZZONE: Quick question.

CO-CHAIR SEPTIMUS: Yes.

MEMBER ARDIZZONE: The injury level that they reported here, that's consistent with what all those other measures are? Or are you creating a new scale?

MR. STEWART: The one we use, none, minor, moderate --

MEMBER ARDIZZONE: Yes.

MR. STEWART: -- I forget, is the NDNQI.

MEMBER ARDIZZONE: Good. Okay. Thank you.

CO-CHAIR SEPTIMUS: Okay. Well, seeing none, we can vote on reliability.

MS. QUINNONEZ: Voting is now open for the reliability of Measure 3003. Option 1, high; option 2, moderate; option 3, low; and option 4, insufficient.

Looking for two more votes. One. All votes are in, and voting is now closed. For the reliability of Measure 3003, 53 percent voted high, 47 percent voted moderate, 0 percent for low, and 0 percent insufficient.

CO-CHAIR SEPTIMUS: Okay. By the way, is Michelle still on the phone?

MEMBER SCHREIBER: Yes, I am.

CO-CHAIR SEPTIMUS: Michelle, you are fantastic. It really takes a lot of discipline to be on the phone this long, so I just want to acknowledge your presence.

MEMBER SCHREIBER: Thank you very much.

CO-CHAIR SEPTIMUS: And if you have any comments, I guess you can raise your hand, and Drew can let us know, okay? Because sometimes I know --

MEMBER SCHREIBER: Okay.

CO-CHAIR SEPTIMUS: But I just wanted

to make sure you were on the line so that you have an opportunity to speak also. All right.

MEMBER SCHREIBER: Yes. Thank you. I appreciate that.

CO-CHAIR SEPTIMUS: Sure. Validity.

MEMBER QUIGLEY: Valid -- oh, so sorry. Thank you. For validity, the same method of validity was utilized as the prior PACE measures, and that is that they had experts to be able to do a face validity of the measures. And then face validity, they had 100 percent agreement on the numerator, and 0.9 -- 90 percent agreement on the numerator. There was no threats tested to validity, so the validity was considered to be high.

CO-CHAIR SEPTIMUS: I want to say it to remind me. This is face validity, so high is not --

MEMBER QUIGLEY: Oh, thank you for that correction.

CO-CHAIR SEPTIMUS: -- is not an option. And --

PARTICIPANT: But 1 is.

CO-CHAIR SEPTIMUS: Yes. So, just -- well first we'll see if there are any comments, but the scale has been changed by our wonderful folks. So 1 is moderate rather than high, just to -- just so people don't get confused when they vote.

But before we vote, are there any comments on this? I guess people are getting kind of used to this stuff here, so let's go vote.

MS. QUINNONEZ: Voting is now open for the validity of Measure 3003. Option 1 of 3 is moderate. Option 2 is low, and option 3 is insufficient. Option 1 moderate, option 2 low, and option 3 insufficient.

All votes are in, and voting is now closed. For the validity of Measure 3003, 84 percent voted moderate, 16 percent voted low, and 0 percent for insufficient.

CO-CHAIR SEPTIMUS: Okay. Next is feasibility.

MEMBER QUIGLEY: Thank you. There were no issues surrounding feasibility. This is something that has to be reported. It's under Type

2 quality reporting. And it's easier to collect data on fall injuries than it is fall rates, so. Thank you.

CO-CHAIR SEPTIMUS: Any comments? Okay. Let's vote on feasibility.

MS. QUINNONEZ: Voting is now open for feasibility of Measure 3003. Option 1 is high, option 2 is moderate, option 3 is low, and option 4, insufficient.

All votes are in, and voting is now closed. For the feasibility of Measure 3003, 32 percent voted high, 58 percent voted moderate, 11 percent voted low, and 0 for insufficient.

CO-CHAIR SEPTIMUS: Okay. Now we're going to usability. And again, just because this is a new measure, remember there hasn't been anything that's publicly been reported, but there is, I think, a plan that this will be publicly reported as part of accountability through CMS. So Pat?

MEMBER QUIGLEY: The issues surrounding usability, obviously, this is going to

improve patient safety. Anything that can be done to reduce injurious falls would improve function, quality of life and reduce mortality.

And also, in discussing usability, hopefully as this goes forward and they continue to -- CMS continues to move forward, is that they really will look at injury reduction strategies and not just base this all on fall prevention. Because there is a body of knowledge in a large health care, national health care system that has actually demonstrated the reduction in injurious falls across settings of care.

CO-CHAIR SEPTIMUS: Any comments? And we'll vote on usability and then we'll go to whether the measure is acceptable for NQF endorsement. So usability.

MS. QUINNONEZ: Voting is now open for the usability and use of Measure 3003. Option 1 is high; option 2, moderate; option 3, low; and option 4, insufficient information.

All votes are in. Voting is now closed. For the usability and use of Measure 3003, 32

percent voted high, 53 percent voted moderate, 16 percent voted low, and 0 percent voted insufficient information.

CO-CHAIR SEPTIMUS: Excellent. Now, suitability for endorsement.

MEMBER QUIGLEY: May I make one more comment please, sir?

CO-CHAIR SEPTIMUS: Absolutely, Pat.

MEMBER QUIGLEY: There was public comments reported for this measure. There were not public comments for all measures, but there was for fall injury reduction. And the comments that came forward were to help CMS and the PACE program to also look at the interface between the participant in their home setting and the use of safe patient handling and movement, in trying to reduce injurious falls associated with safe handling, assisted transfers, assisted mobility, because those patients can fall even with lift devices.

So there were multiple comments in relationship to that, which I think is really

helpful, that would integrate safe patient handling and movement with fall injury prevention, that there is a body of knowledge with, as well, especially in the Department of Veterans Affairs. Thank you. That concludes my remarks.

MS. QUINNONEZ: If there are no other questions, voting is now open for the overall suitability for endorsement for Measure 3003. Option number 1 is yes; option number 2 is no.

MR. STEWART: We're voting. We're going to try to come back at 4:15.

MS. QUINNONEZ: All votes are in, and voting is now closed. For the overall suitability for endorsement for Measure 3003, 95 percent voted yes, and 5 percent voted no.

CO-CHAIR SEPTIMUS: I really want to thank the developers for hanging in there. I know the first measure was a bit long, but I think very fruitful in terms of constructive feedback. The next two measures obviously went much easier, but we really hope that this was instructive for you to take back to maybe revise that first measure that

we did not find suitable.

But thank you for your efforts, and thank you for hanging in there.

MEMBER SMIRZ: Thank you. We appreciate all of the comments of the committee.

MR. STEWART: Thank you for your time and expertise, and especially the meaningful input.

CO-CHAIR SEPTIMUS: So we'll stand adjourned until 4:15. Then we'll talk about opiates. Yes. We've had enough pain, right?

(Whereupon, the above-entitled matter went off the record at 4:03 p.m. and resumed at 4:17 p.m.)

CO-CHAIR THRAEN: All right, we're going to get started again. So this is Measure Number 2940, Use of Opioids at High Doses in Persons without Cancer. We have the Pharmacy Quality Alliance here to present. And Leslie is the lead. So we'll start out with the measure developers.

DR. EISENBERG: Good afternoon, and thank you for considering our measure. My name is Woody Eisenberg. I'm the Senior Vice President for

Performance Measurement at PQA. And I'm joined here this afternoon by Lynn Pezzullo, who is our Senior Director for Performance Measurement, and also by Kristen Butterfield, who is our Director of Research and Analytics, and also on the phone by Lisa Hines, who is a Director for Performance Measurement.

Hi Lisa. Are you there?

MS. HINES: Hi there. Can you hear me?

DR. EISENBERG: Lisa is here. We hear you. Thank you.

MS. HINES: Thank you.

DR. EISENBERG: We have three measures of potential opioid over-utilization that are related, but each one's a little different. They're being introduced to you as three separate measures, and we're going to concentrate, initially, on 2940, which is Use of Opioids at High Dose in Persons without Cancer.

The description of this measure is the proportion of individuals without cancer receiving a daily dosage of opioids greater than 120 mg

morphine equivalent dose, MED, which by the way, is the same as MME, which is what the CDC uses, morphine milligrams equivalent, so those are the same terms, for 90 consecutive days or longer.

A brief background, abuse and overdose of prescription drugs is a major problem in public health in the United States.

(Off microphone comments.)

DR. EISENBERG: Okay. So let's get a little deeper into the measure, then. So as you know, there's no FDA maximum dose or duration for any of the opioid drugs. And studies, though, have demonstrated that patient populations taking high opioid doses for prolonged periods, are often characterized by high rates of psychiatric and substance -- psychiatric illness, substance abuse disorders, and they have high, higher incidences of drug overdoses, and higher death rates.

In 2010, the Washington State Agency Medical Directors Group suggested 120 mg MED as a dosage level that should not be exceeded without special consideration. Subsequently, the Group

Health Cooperative implemented this guidance, and demonstrated a reduction in their opioid deaths for their patients with chronic pain.

Similarly, CMS Part D has adapted these guidelines, the Washington State guidelines, for their over-utilization monitoring system, which all Part D plans implement today, and for the last four years now, to initiate conversations with their prescribers and approximately 40 million Medicare members.

Since the introduction of this system in 2013, CMS has recorded an approximate 25 percent reduction, compared to the 2011 baseline, in total beneficiaries with at least 90 consecutive days greater than 120 mg MED, and greater than three prescribers, and greater than three pharmacies for the opioid claims.

These, in fact, are the parameters that we'll be discussing for the three measures. The proposed PQA measures mirror these parameters, and -- which have, by the way, been built into the CMS program and are now reported to plans, as part of

their patient safety reports.

Also, the high dose measure, the first one that we'll be considering, is included this year in the Medicaid Adult Core Measure Set, so that plans can choose to report this.

As I've told you, PQA has developed three measures related to prescription opioid abuse. The measures examine the quality of use at the health plan-level. I'd like to make that clear, because I know in the comments there was some focus on prescriber-level information, which is not part of our measures.

They're all the health plan level, and they're all related to high dose of the medications over time, and access to medications through multiple providers. And then the third one is the combination of these two, high dose and multiple providers.

I should add that our development process included health plans and PDMs, and these are the two entities that are impacted by these measures. We also included prescribers as

consultants, so we've had input from the physician community as well.

The measure, which is the first one, Use of Opioids at High Dose in Persons without Cancer, focuses specifically on the use of opioids at high dose, and -- thank you.

MEMBER SCHULTZ: Thank you. Okay. Measure 2940 is a new measure. It's one of three related measures, except they are separate measures. It's a process measure. And I think the developer has given a nice overview.

I think, in terms of context, and relevancy, we have an opioid epidemic. We have a new law from our president which mandates education, prevention, treatment and rehab. And with this measure, we would have something to measure. You manage what you measure, and absent a measure, you don't know how you're doing.

So, much as last year, we had a antimicrobial stewardship and use, this year we have opioid use. And I'm sure this group will want to jump right into appropriateness soon enough, but

we have to just get a handle on a national metric for variation and benchmarking, and get -- we have to start somewhere. And this is a wonderful place to start. So thank you for this measure.

I'll start with, it is again, a process measure. It is claims-based, so it's administrative data. It is at the health plan or population level. And as we're calling for more and more accountability for populations of health, it fits in constellation of managing a population to make them safer.

And it also has an aspect of helping us to identify and eliminate waste, so over-utilization of an unnecessary resource, perhaps. So once we start to manage it -- once we start to measure it, we'll get a better handle on that.

So if we start with the evidence, they actually did a very nice job. There is a systematic review of the evidence specific to high doses of opioids and opioids for a long duration.

As I said, there are no proven benefits

for extended and high doses of opioids. Opioids are wonderful drugs, used appropriately. Used inappropriately, or just it's easier or it's quicker to give you a prescription is not going to cut it.

And so this will help for -- although it's at the plan level, the plan will give the feedback to the providers. I'm sure that's how this logic is going to go help me understand why you are way out of line with your counterparts in this particular plan, and perhaps you should be depaneled from our plan. And that's not necessarily a bad thing.

So in terms of the evidence, what they put was contemporary. And it was sound evidence, and so our NQF staffers pre-rated it as high. And I would agree with that.

CO-CHAIR THRAEN: Thank you. Questions or comments about the evidence? Missy.

MEMBER DANFORTH: Can you just briefly describe -- you provided a lot of evidence about the dosage, the 120 mg, but what about the 90 days,

instead of, for example, at 60?

DR. EISENBERG: There is no right number for the number of days. But almost all of the literature on chronic pain uses 90 days. Chronic utilization of opioids is -- I wouldn't say it's defined as 90 days, because that's over-stating, but that's where all of the research has been done, at 90 days.

CO-CHAIR THRAEN: Yanling?

MEMBER YU: Yes, thank you. My question -- I'd really like the measure to address this national crisis, this opioid -- since it really harms lots of people, and it really doesn't demonstrate worth at high dose.

My question is, the goal to have those measure is really to help to improve the prescribing behavior, and therefore improve the safety and the quality. The level of analyses is at the plan level, so Medicare, Medicaid and a commercial health plan. So I was just wondering, have you thought about why this not at a facility level and is more directly connected with the prescribing

pattern and the communication between provider, who really prescribes the drug, and the patient, who receives those?

The improvement could be at a facility level, rather than at the plan level. That's -- I wonder you have any thoughts on that.

DR. EISENBERG: Yes. That's an excellent comment. Thank you. To begin with, we want to start out using the tools that we have. And we have a wonderful tool in Medicare Part D, which has 40 million Medicare members in it. And that's called the Stars Rating System.

The Stars Rating System consists of feedback to the plans and ratings that eventually impact how popular they are in terms of their choice and payment.

Our measures today are part of a patient safety reporting system that CMS uses to give feedback to the plans. The plans then take that feedback, and they then contact their prescribers, their pharmacies, their members. This is part of the over-utilization monitoring system.

So that's already in place. So we thought it was best to start there, to use the tool that's in place. And similarly, in Medicaid, as that, as the core adult set expands, and becomes a more leverageable tool, we think that it'll work well there, also.

But we're not satisfied, and we don't intend to stop there. We're in discussions right now with CMS, to develop patient/prescriber-level measures that are based on our measures that we're discussing today.

We're also in discussion with other measure developers, NCQA, to have this, or a version of it, added to HEDIS, so that it'll impact commercial plans. We think there's lots of different areas that we, where we can go. This is where we're starting.

MEMBER YU: Okay.

CO-CHAIR THRAEN: Kimberly?

MEMBER APPLEGATE: I agree that it's a really important measure. I had a question about the -- to the developers about understanding

missing data, and if there was any consideration of what's been in the news about when the patients can't get the prescriptions filled, if there's too much attention or unintended consequences on too much attention to not filling scrips or decreasing script delivery, that family, friends, street sales will go up, and we won't capture that. So I just wanted to address that.

DR. EISENBERG: The measure is -- let me address your first question first. The measure is based on administrative claims data, which has really shown virtually no missing data. And in order for plans to get paid by CMS for the services they provide, they have to provide all completed claims to CMS.

Those claims are scrubbed first by the individual health plans. They are then sent over to CMS, where their contractor further scrubs them to turn them into what's called prescription drug events.

Everybody works real hard to get every piece of that data. It's a natural -- and the

measures are just a natural byproduct of that, so we are very confident that we have virtually all the data.

In terms of the second point you raised, I think this is a very important consideration. As a society, we're always wrestling between under-treating patients in pain, and over-treating patients who may legitimately have pain or not. And I think that pendulum swings back and forth, sometimes to extremes.

So right now, as a society, we're really focusing almost entirely on over-utilization of these drugs. And yes, we are concerned that they're -- that this could be part of an effort in this country that pushes things a little too far so that some patients may in fact be under-treated.

But I would just add that, as you know, virtually every agency in Health and Human Services has a program now, for this. So this would be one small part of one program.

MEMBER APPLEGATE: My only other comment, just for the record, just for the record,

is that we're not providing behavioral services in psychology and psychiatry to back up what we're trying to do. And so I think we may end up with unintended consequences.

CO-CHAIR SEPTIMUS: Just to remind people, we still have to go through the evidence, and we haven't voted on that yet, and you're already jumping to measures and stuff. So let's go in order.

CO-CHAIR THRAEN: Steve. No? Are there any other questions or comments, just related to the evidence. Steve?

MEMBER LAWLESS: Do you sort out by state? Some states have very, very strict requirements of documentation of opioids, treatment plan, very detailed, and if you don't, there are penalties. If there's grievances, I mean, people will, could lose their licenses, provision.

Sorting out that has the biggest impact in terms of reducing people who are on medications, because the onerous requirements to document why

they're on it versus just having a measure. Do you -- are you able to sort, state by state, those differences as that impact?

DR. EISENBERG: We could. The data comes in depending upon the program, of course. If it's a Medicare program, then the data will be national, but it will be parsed according to the health plan.

So if there's a health plan that works only in New York State, then that's the information we'll get from that health plan. But there may be another health plan that works in New York, New Jersey and Connecticut, and there, the data will be at the contract level, meaning no, was the answer to your question.

MEMBER LAWLESS: Okay.

DR. EISENBERG: This year, for the first time, Medicaid is implementing the high dose measure. That will clearly be state-specific. And then as this moves out into commercial areas, we anticipate that it could even be down to regional and perhaps even cities.

CO-CHAIR THRAEN: Any other questions about the evidence before we vote? All right. Let's vote.

MS. QUINNONEZ: We are now voting on Measure 2940, Use of Opioids at High Dosage in Persons without Cancer. 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 2940, 70 percent voted high, 30 percent voted moderate, 0 percent for low, and 0 percent for insufficient.

CO-CHAIR THRAEN: All right. Leslie, you want to cover performance gaps?

MEMBER APPLEGATE: Okay, this measure was tested in three different health plan sources: a Medicare population, a commercial plan and a Medicaid population.

The testing in the Medicare population was a huge number, 7-some million, I believe. Medicaid range was over a million, and the

commercial plan was perhaps the smallest. It was only about 200,000 patients.

There is distribution, and there's variation across the plans, and then definitely within the plans. And this measure also has evidence of disparities, in terms of a lower socioeconomic, in terms of people who are getting the low income subsidy.

Their measure of use was the greatest. It was 62.4 per 1,000, which is like double, triple the other patient populations. So it's like, who are these people, and is it just quicker and easier to give them a script and just be done with it?

So there's a great performance gap in variation, so opportunity for improvement.

CO-CHAIR THRAEN: Questions? Yanling?

MEMBER YU: In the gap estimate, I can see, definitely there's a big gap. It'll be for Medicare and Medicaid. Now, this commercial health plan, I'm sure maybe you have data, but it's not sure. You didn't show 25 percentile, 50, 75 and

quartile range. You only quoted mean. Is there any reason for that?

MS. BUTTERFIELD: Yes. The reason is because we only had access to one commercial plan, so there was no way to, in our testing, look at a distribution, because there was only one plan that was included in the analysis.

MEMBER YU: Okay. So there will be just one?

MS. BUTTERFIELD: Just one commercial plan.

MEMBER APPLEGATE: And a level of measurements at the plan level.

MS. BUTTERFIELD: Correct.

MEMBER YU: Okay. Yes.

MS. BUTTERFIELD: Yes.

MEMBER YU: Okay, just one. Okay. Oh to -- yes, you mentioned one plan. I'm sorry.

CO-CHAIR THRAEN: Questions? Shall we vote?

MS. QUINNONEZ: We are -- voting is now open for performance gaps of Measure 2940. Option

number one is high, option number two is moderate, option number three is low, and option number four, insufficient.

Here it is. That should be 19. Voting is now closed. All votes are in. For performance gaps, we have 84 percent voted high, 16 percent voted moderate, 0 percent for low, and 0 percent for insufficient.

CO-CHAIR THRAEN: All right, reliability.

MEMBER APPLEGATE: Okay. For reliability, it was a signal-to-noise analysis, who looked across the three different groups, the commercial, the Medicare and the Medicaid.

For the Medicare testing, it was a sample, a convenient sample of over 700 Part D plans, compromising a total of over 7 million patients, aged 18 and over.

The commercial plan, again, had the smallest number, but it was one plan. And the Medicaid testing, it included eight state-based prescription drug plans covering six states, and

again, there were 1.4-plus million patients included as, in part of the sample.

The reliability scores, the mean reliability was rather impressive, nearing 1. The minimum was a 0.98 and the maximum was a 0.99, so incredibly high signal-to-noise ratio here.

MS. BUTTERFIELD: Can I speak to that very quickly?

CO-CHAIR THRAEN: Sure.

MS. BUTTERFIELD: We actually only did the reliability testing within the Medicaid population, so just to state that. So there was eight plans, or eight -- yeah, eight plans from six states that were included in the reliability analysis, and that was mainly because of data access, as well that the measures being used in the Medicaid adult core set. So that's where the reliability statistics came from, was the Medicaid population.

MEMBER WU: Was that because --

MS. BUTTERFIELD: And not the Medicare or commercial one.

MEMBER WU: Was that because you couldn't get a drug-dependent patients out of Medicare, or what was the problem?

MS. BUTTERFIELD: To do reliability testing, you need plan level data, and we were not able to get contract level data for Medicare. But we had that for Medicaid, and that's where we -- that's why we did the reliability testing within the Medicaid population.

MEMBER APPLEGATE: Thank you for coming forward with that.

CO-CHAIR THRAEN: Lisa?

MEMBER MCGIFFERT: I -- can you explain the denominator again to me? I'm having some trouble with some of the days, supply is greater than or equal to 15, and I don't know, I'm missing something here.

MS. PEZZULLO: Right. So the denominator looks for individuals who received prescription -- who had prescription claims for -- two or more prescription claims for opioids, where when you sum the day's supply, there are at least

15 days' supply when you're summing those two, or possibly more.

And the reason being that we wanted to focus on those individuals that are potentially using these for chronic use, and eliminate those smaller -- exactly, you know, a dental procedure where you have a five-day supply type of thing. So thank you.

CO-CHAIR THRAEN: Yanling? And then Missy, you had yours up, and then Jason.

MEMBER DANFORTH: I just want to clarify -- so Lisa asked my question, but now I have a clarifying question. So if you're trying to eliminate the sort of one-time use, the 15 days makes sense, but why the two separate prescriptions?

What about a -- because what you'd be missing is basically like a one-time prescription for 90 days. I mean, you could be missing what you're looking for on the numerator by virtue of how you've defined the denominator.

MS. PEZZULLO: I -- you know, just based

on the requirements and regulations around dispensing of these types of medications, it's not likely that they would be dispensed as a 90-day quantity, 90-day supply. It's more likely that the maximum would be a 30-day supply, in most states.

CO-CHAIR THRAEN: So -- okay. Hold on. So Kendall, you said something related to that?

MEMBER WEBB: Sorry. I'm actually out of order. You got me out of order. But, I mean, you're still not going to catch one if it's a 30-day supply. I can tell you, my ortho docs routinely prescribe 30-day, and then if they can, 90 days. You can't really do 90 days anymore until you've already done your 30 days, but single prescription, they'll get as many out of it as they can.

CO-CHAIR THRAEN: I think it's been duly noted. Jason, and Yanling, did you still have a question? Okay. Jason first, and then I'll come back to you.

MEMBER ADELMAN: I have two questions. I'm sorry if I missed it, but did you say what was the intended use of the measure, meaning like, is

it to monitor particular physicians and their practice, or is it whole health plans, or all of the above? How do you intend to use the measure?

DR. EISENBERG: The measures are all at the health plan level. Now the health plans have great incentive to be in communication with their physicians as well as with their patients, but the health plans at the contract level are what are actually being measured.

MEMBER ADELMAN: And my second question was again about the denominator. I understand, I understand the denominator, I just don't understand the rationale for the denominator, meaning like, for example, what if it was, instead of what it is now, if it was all patients? What is the benefit of this versus all patients?

It would still -- you said something about excluding those that have, only take it for a few days, but the numerator does that by the nature of what the numerator is. You have to be on it for 90 days straight at a very high dose. So what is the benefit of the current denominator versus all

patients?

MS. PEZZULLO: Right. So the discussion within the measure development workgroup was to, if you include all patients, then you could potentially be inflating your denominator inappropriately, because they wouldn't be included, potentially, in the numerator.

MEMBER ADELMAN: But why is that inflating the denominator or just accurately reflecting the denominator, meaning like, if you have a hundred patients with chronic meds, over a thousand, or a hundred over a million, then you're doing much worse if you have a hundred over a thousand.

And just because you have a million patients doesn't mean you're inflating it, just means you're taking care of a lot of people and you're much bigger. So I don't follow the logic. And the denominator confuses me.

And I feel like I, if I think about it, I could feel like that's introducing a bias, because now the denominator is people that have lots of

patients on some kind of narcotics, for whatever reason, and so I don't see the logic for it, and I see the potential for introducing a bias.

MS. PEZZULLO: Yeah. So I'll go back to just the decisions made by the measure development workgroup was, the definition of the denominator was more an attempt to focus on those that are more -- a denominator that better defines chronic use, and eliminating some of those just very short-term, acute type of prescriptions. But I hear your point.

CO-CHAIR THRAEN: Yanling?

MEMBER YU: Yes, thanks. My question is about the reliability of implementing the measure for the commercial plan. I know -- I understand Medicare and Medicaid would have, you know, lots of drive to really, to adopt this type measure, but what about the commercial plan? Do you have a -- health plan, do you have any thoughts to share like, something like that?

You know, you -- when you did the gap analysis, you only had one plan, really didn't have

two or three plans and to really do a comparison. Now I'm just wondering, you know, what would drive them to adopt it?

DR. EISENBERG: Yeah, right now there really isn't a lever, other than public pressure, the knowledge that they all have, which we all have, that this is an epidemic, and the fact that the prescribers that are in commercial plans are also in Medicare and Medicaid plans, so it's likely that there will be trickle down to commercial plans as well. But there isn't anything to force them to adopt this right now.

CO-CHAIR THRAEN: There is a cultural shift going on, towards value-based payment. And the -- Medicaid and Medicare are moving towards accountable care organizations, which are managed care versions of the local health plans.

And as they start to implement some of these things for the Medicaid and the Medicare population, it will also bleed into the commercial plans, the non-Medicare and the non-Medicaid plans as well. So it's moving in that direction. Laura

and then Albert.

MEMBER ARDIZZONE: Thank you. About your exclusion criteria, I absolutely understand the exclusion of cancer and patients in hospice, coming from a cancer institution. But I was also thinking about patients with chronic conditions, so cystic fibrosis, sickle cell, HIV.

Has there been any thought of maybe not excluding them, but stratifying for those? Maybe the Medicaid plans or something like that may have a higher proportion of patients with some of those chronic illnesses, who are a little different than opioid-seeking patients.

DR. EISENBERG: Yes. There was lots of conversation, and we've also received lots of consults who have a variety of different opinions about that.

I think our final decision not to exclude those patients had to do with the evidence from the CDC guideline, among other places, that was not able to demonstrate efficacy of higher doses, so that although these patients may have chronic

pain, they may have severe chronic pain, there doesn't seem to be evidence to support that they need a dose higher than 120 mg, that it benefits them. In fact it may be worse for them.

So we have actually had passionate arguments that these patients should clearly be included because it's for their own good. And some of our consultants that also participated in the CDC effort, have been now addressing, for example, the cancer patients that are five-year survivors and trying to get their lives together, but are addicted to opioids, so.

CO-CHAIR THRAEN: Albert, did you have?

MEMBER WU: Just going to comment that commercial plans, particularly disability insurers, can exert influence on their, who they cover, and they can, in fact, refuse payments, or send messages to providers who routinely prescribe higher than needed doses, and we've actually worked with some private plans that have managed to significantly lower their prescriptions of opiates, just by doing that.

CO-CHAIR THRAEN: I think, in reality, the environment's looking for something. It's been looking for some definitive recommendations to move in this direction. I think the plans are actually going to embrace it quite well.

Shall we take a vote on the reliability question? Did I miss somebody? Missy, yeah?

MEMBER DANFORTH: So if we disagree with the denominator, is -- that would impact the reliability voting? Because I think Jason disagreed -- had some problems with the denominator. I have some problems with the denominator. So --

MR. LYZENGA: I might argue that it's the validity of the -- I would sort of see reliability as a question of whether the denominator and other specifications are clearly and precisely defined. And then if you think that they are not actually reflective of the evidence or the, you know, of quality, then that would go under validity. Does that make sense?

CO-CHAIR THRAEN: And just so that I

understand and everybody else understands, what's the disagreement with the denominator again, just quick, you know, in a reframe?

MEMBER DANFORTH: It seems like the way it's defined, which is two or more prescriptions ---

CO-CHAIR THRAEN: Oh, okay. It's --

MEMBER DANFORTH: -- on at least two separate dates, like you're missing a big --

CO-CHAIR THRAEN: Okay.

MEMBER DANFORTH: Yeah.

CO-CHAIR THRAEN: Call for the vote?

MS. QUINNONEZ: Voting is now --

MEMBER WEBB: Can I just put one more thing in? What do we do with trauma centers? What do we do with trauma centers? This is going to kill trauma centers. What we're going to do --

CO-CHAIR THRAEN: So the question is, what do we do with trauma centers? I'll give that to the developers.

DR. EISENBERG: Could you elaborate, why will this kill trauma centers?

MEMBER WEBB: Because we see a lot of

patients who require more than two prescriptions, even in a 30-day period sometimes. I mean, most of the trauma centers are visited by the low-income population. It's a lot of disabled patients.

I think this is a place where your socio-economic -- you're just not going to get the same -- you can't compare inner-city urban trauma centers to a community with a payer mix that is, you know, primarily private.

PARTICIPANT: This is not a hospital case, though. This is a plan case, right, which --

CO-CHAIR THRAEN: So --

PARTICIPANT: So outpatient.

CO-CHAIR THRAEN: Your clients have no plan.

MEMBER WEBB: Okay.

CO-CHAIR THRAEN: Yeah. They will have a Medicaid coverage or no coverage, for you, likely, for the uninsured. Basically the folks that you're talking about, socioeconomic folks fall into that category. This is a measurement at the plan level.

So this, the plans who cover a variety of folks that show up at your trauma center, they're rolled into all the other ones that are not showing up at the trauma center, the measures.

MEMBER WEBB: So, like my payer mix is 45, 50 percent Medicare, Medicaid, more Medicaid than Medicare. My Medicaid plans potentially are going to get penalized, correct?

CO-CHAIR THRAEN: Only for your behavior. But the trauma, I think the trauma question is going to -- is a generalizable question across plans. It's not going to penalize a specific plan.

MEMBER MCGIFFERT: Is it outpatient or inpatient? Or is it every prescription?

DR. EISENBERG: This is outpatient. And it -- right. Ninety days -- yeah.

MEMBER DANFORTH: Wait, but the denominator is still two prescriptions -- 15 mg, outpatients, okay.

CO-CHAIR THRAEN: Michelle?

MEMBER SCHREIBER: Thank you. You

know, I have to agree with the concerns about the trauma center. We're an inner-city trauma center, and our inpatients and our outpatients might fall under the category of patients who are using this more often.

You made the comment of a higher use in urban areas, and that it may be just because we're writing scripts because it's easier. That's not true. We have a disadvantaged population who have had gunshot wounds, chronic pain, HIV, sickle cell, spinal cord injuries. And I really fear that some of these plans, such as our plan, would be penalized.

Our patients are in plans. Medicaid patients in the State of Michigan all have to be in some kind of a plan. Plus we have the dual eligibles that are in some kind of a plan.

So I guess I'm looking to see what your plans are, I guess, for either stratification or risk adjustment for some of these issues, or your justification for not doing it. Thanks.

DR. EISENBERG: The way the data has

been analyzed and presented to this committee is that it would be stratified differently for Medicare, Medicaid and commercial, because we noticed, as you have, that there are greatly varying degrees of measure rates, two to three times for Medicaid plans, for an example. So yes, that is part of the plan.

CO-CHAIR THRAEN: Okay.

MR. LYZENGA: Again, I would just say that your voting on that particular issue should probably be reflected in the validity as well.

CO-CHAIR THRAEN: Okay. So we're back to voting on reliability.

MS. QUINNONEZ: Voting is now open for the reliability of Measure 2940. Option one, high, option two, moderate, option three, low, and option four, insufficient.

All votes are in, and voting is now closed. For reliability of Measure 2940, 62 percent voted high, 33 percent voted moderate, five percent voted low, and zero percent for insufficient.

CO-CHAIR THRAEN: All right, validity now.

MEMBER APPLEGATE: Okay. So with this new measure, what they have provided us is, at the measure score, face validity only, which is a limitation for us, in terms of, we now cannot choose high as an option. We have to start at moderate.

It was an expert panel, and there's some questions and some comments as to the composition, and was that expert panel representative, because they're -- in part, they're part of the development process of the measure.

However, they did seem to be pretty representative of industry and pharmacists, in general, so I took great comfort into sort of the array of the experts who were on the panel, and their credentials.

So in terms of the ability for this -- does this variable vary, it does vary. And then the analyses were conducted amongst the Medicare population and the Medicaid population, and there are distributions of the performance.

There's a little less variation amongst the Medicare population. And since we don't have a frame of reference, I don't know if these numbers are -- strike me as just like too much to begin with, even though there's not a lot of variation.

But definitely amongst the Medicaid population, there's a vast amount of distribution from a minimum of 8.15 to a maximum of 6.645, which seems like a lot of variation. So, you know, until we have the measure and we have national experience, we really don't know is -- what's the right number.

Right now, we've got a lot of variation, and it looks like a lot of opportunities to understand the next level of drill-down, but we need the measure first. So.

CO-CHAIR THRAEN: Questions? Laura?

MEMBER ARDIZZONE: You know, I was -- this is for the developers. I was -- I didn't understand NQS' initial comments that some of the faces -- face validity was done by the same experts and stakeholders. Because when I cross-referenced the list, it looked like there was only one person

who was the expert panel who was also on the measurement development list. Is that correct?

MS. PEZZULLO: Yes. So your interpretation is correct. So they were two separate groups. So we had individuals that participated on what we call the measure development workgroup, that really gets -- digs in and defines all the aspects of the measure. And then our quality metrics expert panel is a different group of individuals that assess the measure specifications once they've been forwarded along by the measure development group.

CO-CHAIR THRAEN: Yanling? I can't talk anymore. Yanling. There you go.

MEMBER YU: Perfect. Thank you. Quick question. On this meaningful difference, for Medicare population, there's no P value coded for the inter-quartile range, but there's one for Medicaid. So do you have the number, just happened, do you have?

MS. BUTTERFIELD: Again, that had to do with data availability, so we were able to do that

for the Medicaid plans, because we had access to plan-level data, whereas the Medicare data, we just had information on the distribution, but we didn't have plan-specific.

There's 700-plus contracts with the Medicare population, and we had the overall data but not data for each and every separate 700 plans, if that makes sense. So we weren't able to do P-value testing based on that, but we do have the percentiles and the inter-quartile range, and the standard deviation for that population. For Medicare it was -- do I have it here? It was 8.32, which shows there's some variation there.

MEMBER YU: Okay. Thanks.

CO-CHAIR THRAEN: Michelle, I think you wanted to make a comment about validity.

MEMBER SCHREIBER: No. My comment was from before, and it was about the question, not just the stratifying, I guess, by Medicare and Medicaid but by socioeconomic and demographic stratification.

CO-CHAIR THRAEN: Okay. Did you get

your comment made?

MEMBER SCHREIBER: More or less, I think. Thanks.

CO-CHAIR THRAEN: Okay. Any other questions or comments on validity? Shall we vote?

MS. QUINNONEZ: Voting is now open for the validity of Measure 2940. Option one, moderate, option two, low, option three, insufficient. Option one moderate, option two low, and option three insufficient.

All votes are in and voting is now closed. For the validity of Measure 2940, 67 percent voted moderate, 33 percent voted low, and zero percent insufficient.

CO-CHAIR THRAEN: All right, next one's feasibility.

MEMBER APPLEGATE: These are claims data. These data are being collected currently. They're used in other programs right for now, so they're probably pretty solid and they're probably pretty clean, given what you described, in terms of the process.

CO-CHAIR THRAEN: So I just want to ask a quick question. So one of the things that claims data that often doesn't give you, is the specifics. So in claims data for prescription drugs, you're getting dosage, you're getting frequency, you're getting all of the above. So you -- it's a really easy analysis. I mean, not easy, but, you know, okay. It can be done, easily.

DR. EISENBERG: Yes.

CO-CHAIR THRAEN: Yanling, and then Lisa.

MEMBER YU: Okay. My question, I think, is about the -- on page nine, it said, a certain use of measures are only approved by license agreement with the -- from the development, and that you were involved some of -- you have -- you were going to reserve the right to determine the condition under which were approved or licensing fee may be even charged.

So could you explain to me that, what do you mean, certain use of the measure? And what do you envision that -- what kind of things that you

will be charge? Will that be a, encourage them to use for a certain way, or it will be hamper them, make them don't want to use your measure? So which way to go? How do you evaluate that?

DR. EISENBERG: The measures are free for use for all of the federal and state programs.

MEMBER YU: Right.

DR. EISENBERG: There's a whole industry that's grown up around advising the health plans, calculating measures for them, tutoring them on how to do the measures, doing the calculations for them.

These businesses that have grown up around the measurement need to have the measures and current NDC lists. They're the ones that we ask to license the measures.

MEMBER YU: So it's not really the plan itself, whether they adopt your measures. That -- I'm trying to understand the fee. So it's not if someone's a plan, particular plan said that, I want to use, adopt your measure, or implement, then you will charge them fee for doing that?

DR. EISENBERG: No. And in --

MEMBER YU: That's not the --

DR. EISENBERG: -- in fact, let -- yeah. It's -- it varies by program. So I'll give you the biggest example, Medicare. The plans don't actually do the calculations. The plans don't actually need the measures.

All the plans do is submit their claims data to CMS, and the CMS contractor does all of the calculation. The contractor and CMS are not charged a licensing fee.

MEMBER YU: Okay. Will you do for free, just let them to use it and, you know --

DR. EISENBERG: Well yes, to --

MEMBER YU: -- to encourage more, you know, commercial plan to adopt this, you know, to improve the -- to encourage more, you know, wide adaptation of this measure?

DR. EISENBERG: For the most part, yes. And it's not just plans. It's also state alliances, state departments of health. There's lots of organizations that we work with. And for

the most part, there's no licensing fee. It's the organizations, the companies that are out there making a profit, in order to be vendors to these various plans, that we ask for licensing.

MEMBER YU: Okay. Thank you.

CO-CHAIR THRAEN: Lisa, then Missy. Missy?

MEMBER DANFORTH: Just quickly, I know you only tested this with one commercial plan, so I think it's important to understand if they had any feedback, or if their experience with the measure was different than the contractor's.

MS. BUTTERFIELD: We did not get any feedback like that from our commercial plan that was tested.

MEMBER WEBB: Did it -- just to be crystal clear, any commercial plan, like Aetna, Cigna, they can use the measure at no cost, correct?

DR. EISENBERG: So if Aetna and Cigna and the rest of them, if they're -- if they need to use our measures so they can improve their own internal quality performance, there's no charge.

If Aetna and Cigna and the rest of them decide they're going to create a new product, that they're going to be selling to Empire, in New York, we would ask them to license our measure.

MS. PEZZULLO: So just, basically as a general rule of thumb, where there would be a licensing fee involved is where others are using the measures within a, I'll say, quote -- I'll say, a commercial product, not necessarily a commercial plan, commercial product where they are making money from selling their product to others.

So that's where, typically, that's the general place where there would be a licensing fee involved.

CO-CHAIR THRAEN: All right. I'm going to call for the vote.

MS. QUINNONEZ: Voting is now open for the feasibility of Measure 2940. Option one, high, option two, moderate, option three, low, and option four, insufficient.

All votes are in, and voting is now closed. For feasibility of Measure 2940, 60

percent -- 62 percent voted high, 38 percent voted moderate, zero percent for low, and zero percent for insufficient.

CO-CHAIR THRAEN: All right, usability.

MEMBER APPLEGATE: Okay. Currently, this is a new measure. It's not currently being used for public reporting. It is used in accountability for the Medicare D over-utilization monitoring system, so it is being used there, currently.

I think it hasn't been around long enough, so if we're talking about improvement results, this is the initial endorsement, so we have to start somewhere. But ---

CO-CHAIR THRAEN: What's the planned usability?

MEMBER APPLEGATE: Right now it says there is no planned use, however, given the recent law, and Medicare's interest in moving to population health, I would not be shocked or surprised if they put it out there soon in the ACO

world for sure.

CO-CHAIR THRAEN: This is a great example where the departments of health get the -- we have an all-payer database, claim database now, and I can see the departments of health using this measure to evaluate what's going on in their own environments. So I see this as very useful.

Any comments or question -- oh, Laura?

MEMBER ARDIZZONE: Just quickly, just, I have a question. It said, CMS has announced plans to move this measure into 2019, Part D, display measures. What's a display measure?

DR. EISENBERG: Part D has a sort of a tiered performance measurement system. The big deal is the star ratings, and they're public information. The plans have to perform well on them, or they can get tossed from the program. And if they perform really well on them, they get bonus standings. So that's a big deal.

The next tier down is a display measure. A display measure means that there's public information available. You can go to the site and

see what your health plan did for the measures that are display measures.

They're also the basis for compliance actions. That means that CMS has a dialogue with the plan, and if they're not happy that they've corrected a problem, the plans, again, can have restrictions, in terms of how they're marketing their plans, whether they can move into new areas, et cetera, et cetera.

And then the third layer, which is where the measures are now, is in patient safety reports, which are just discussions between CMS and the individual plan, not made public.

CO-CHAIR THRAEN: Okay. Let's call for the vote. Call for the vote.

MS. QUINNONEZ: Voting is now open for the usability and use of Measure 2940. Option number one is high, option number two, moderate, option number three, low, option number four, insufficient information.

Okay. Voting -- all votes are in. Voting is now closed for the usability and use of

Measure 2940. 52 percent voted high, 43 percent voted moderate, five percent voted low, and zero percent voted insufficient information.

CO-CHAIR THRAEN: And then finally, endorsement, suitability for endorsement. Go ahead.

MS. QUINNONEZ: Voting is now open for the overall suitability for endorsement of Measure 2940. Option one, yes, option two, no.

All votes are in and voting is now closed. For the overall suitability for endorsement, 100 percent voted yes.

CO-CHAIR THRAEN: All right, moving forward. So we have two more measures in this cluster. The next one is 2950, Use of Opioids for Multiple Providers in Persons without Cancer. And Laura is the lead.

Do you want to -- do you think you can summarize, kind of, what the specific differences might be, between this one and the one we just did?

DR. EISENBERG: Yes. I can do that --

CO-CHAIR THRAEN: You can do that?

DR. EISENBERG: -- easily.

CO-CHAIR THRAEN: Never mind.

DR. EISENBERG: This one does not rely on dose or duration at all. However, what we're looking now is the proportion of individuals without cancer receiving prescriptions from four or more prescribers, or -- and, and four or more pharmacies during the measurement period.

It's doctor shopping, and pharmacy shopping. And the basis for it is that there is honestly moderate evidence that there's a relationship between numbers of prescribers, numbers of pharmacies and patients having bad outcomes of drug overdose and higher death rates.

Although there's no consistent evidence-based definition of what that means, doctor shopping and pharmacy shopping, several studies have demonstrated that patient populations receiving medications from four or more prescribers and four or more pharmacists have a higher incidence of these bad outcomes.

CO-CHAIR THRAEN: Kendall?

MEMBER WEBB: So, just another question. We have this fight in Florida right now, where our E-FORCSE, which is our ability to see where other opioids have been prescribed, is terrible. And there's nothing we can do it -- do about it.

What do you see as the plan for how to use this in a way to enforce decreased opioid use?

DR. EISENBERG: The present system consists of the prescription drug monitoring programs. I assume that's what you were talking about. There's 50 of them. They're different in every state. They don't talk to one another. There's lot of problems, but they'll get better.

What this measure would do, it was -- it would elevate the responsibility for monitoring the multiple prescribers and multiple pharmacists, to the health plan level. It would place the responsibility on their shoulders. They would then use the tools that they have with their physician and pharmacy networks to take care of the problem.

CO-CHAIR THRAEN: Yanling?

MEMBER YU: I just want to mention the article I just recently read. The article said that -- I found this in a journal. I forgot what was the name. The evidence of PM -- what you called the physician prescribed monitoring program, was it PMT -- does not seem to correlate with the decreased use of opioid medication.

As -- I don't know if you'd say that's a recent article just published by BMJ or I've forgotten.

DR. EISENBERG: So is the question, are prescription drug monitoring programs working to decrease opioid overuse? Is that what you're asking?

MEMBER YU: Well I'm just mainly pointing out that there was an article just recently published then.

CO-CHAIR THRAEN: One of the challenges is that in -- historically, I don't know, you mentioned the other states, I can only speak to Utah, is that the controlled substance databases

have been -- are not even in the Department of Health or in the Department of Human Services where mental health and substance use disorders are located. It's in the Department of Commerce, under the Division of Professional Licensing, and have been a Criminal Justice data source.

And it's only been in the last four to five years where there's been this push from the point of view of health, to say, this is a resource that ought to be integrated into the electronic health record, so that at the point of care, when the prescriber's making the decision, the information is pushed out to the prescriber to say hey, wait a minute, this stuff is on board.

It's a very clunky, very old system. They have to get out of their electronic record, go into the state-based system and wait for the Internet to catch up. And it goes to California for a while, and then it comes back.

And then there's lots of security issues associated with that, because it's located in government, and they have to get through the

firewalls, et cetera, et cetera, et cetera. So current -- the functionality to date has not been a positive effort to support providers in having that knowledge at the point of care when they're making the decisions. That's been our experience in Utah.

MEMBER WEBB: And I concur out of Florida. If you move machines, and forget to log out of the first machine, then you have to create a new password. And you can't ever create the same password.

So if you work anywhere other than a place where you sit at the same machine all day long every day, you -- using E-FORCSE is just unbearable, because it uses the cache of a particular machine. So it just wasn't designed very well.

CO-CHAIR THRAEN: And I want to make one other point here, which is, the traditional approach has been to monitor the provider, as opposed to provide decision support to the provider.

And the new direction is to provide

decision support at the point of care, so that the provider can make the decision as to what is an appropriate prescription, once they have that knowledge, they have access to that information.

So the monitoring approach is an after-the-fact approach. It's not at the point of care.

MEMBER LAWLESS: What's your definition of provider? A multi-specialty group, multi-person group, each individual person? Is it a practice? Is it NPI? What is --

DR. EISENBERG: It's NPI, which means that it's the individual prescriber.

MEMBER LAWLESS: So is that -- so in terms of that, then, in terms of a -- is this trying to drive access to a single provider? Is this trying to -- I -- that seems a little bit of selection out of individual doctors who usually don't have a lot of other people providing a lot of due diligence around what they're doing, versus -- I mean, why NPI and not group?

DR. EISENBERG: You know, I don't know.

MEMBER LAWLESS: What is the practicality?

DR. EISENBERG: I don't know what the practicality is doing.

MEMBER LAWLESS: What do the practice plans do with that, then? Because I have a group of five of us, different people on call, different coverage systems. Automatically, by having a large group, that may select out for this, and may be earmarked.

MS. PEZZULLO: Right. So that is one of the reasons why the workgroup, when they were developing this, wanted to -- so there was some discussion around, should it be receiving prescriptions from four prescribers and four pharmacies, which is the current version, the measure that we're putting forward.

But there was also, you know, should it be four prescribers or four pharmacies, because there can be, you know, concerns with either. But when you -- so to take -- I guess you could look at it as a more conservative approach from the measure

perspective, using and.

So at a minimum, this kind of implies or identifies less than optimal coordination of care. So there could be as, you know, some have discussed earlier, it could be, you know, doctor shopping or pharmacy shopping.

But there is also, you know, when you look at it from the safety aspects, it's -- you know, when it's -- is it likely that somebody might see four different providers even if they're within the same group? Possibly.

Is it likely that -- you know, but when they're getting these opioid prescriptions from four or more prescribers, and four -- getting them filled at four or more pharmacies, that is where there's a greater risk for harm.

CO-CHAIR THRAEN: Lisa?

MEMBER MCGIFFERT: Well I think my comment was going to be in response to something that you were saying about getting to the provider level to help them, but basically this measure is a doctor shopping measure, right?

I mean, this really is getting at who's, you know, who's shopping around for doctors. And I guess, I don't know that it would help at the doctor level, because the -- it's a plan level measure.

So you wouldn't really -- the doctor wouldn't see which patients were shopping around. They would just see that maybe this plan isn't controlling that kind of shopping around, correct?

DR. EISENBERG: Well it's -- this -- these measures are part of a larger system, right. I mean, there's no measure that's going to accomplish the goal entirely. And there will be lots of different systems, but let me describe to you the Medicare system that's in place right now.

Based upon these parameters, the same specifications that are built into these measures, what CMS does is they notify the health plans. They say, this patient and these doctors and these pharmacies are over the limit. And it's your responsibility, health plan, to do something about that.

And CMS has really changed its own rules, in terms of what plans can do, in terms of notifying members, and in terms of what the plans themselves can do to put in patient-level prior authorizations, something they've never done before.

So they're -- it works out as part of a system, not independently.

CO-CHAIR THRAEN: Laura, and then Steve, did you still have a question? Laura.

MEMBER ARDIZZONE: I guess, quickly what I wanted to say is, your decision support for the prescriber or the provider is important, but this, I don't think, is what this measure is trying to address. This is trying to address doctor shopping, multiple prescriptions, multiple pharmacies, multiple providers prescribing.

The state level provider monitoring programs, as we said, do not do enough, are so different in every state. This will elevate this measure to something where we can start making a reliable change, and impact.

And that's why I think, for the first question, if I would be so bold as to say we should move towards voting on the evidence, because I think we agree that we -- this is strong evidence, or moderate evidence of --- there's no systematic review, but moderate evidence, and that this is an important topic.

CO-CHAIR THRAEN: So I would -- I mean, I don't disagree with you said. I would alter it to say that it helps get the physician out of the enabling role.

MEMBER ARDIZZONE: Can we please say provider? Because nurse-practitioners --

CO-CHAIR THRAEN: You're right.

MEMBER ARDIZZONE: -- across the country --

CO-CHAIR THRAEN: Absolutely.

MEMBER ARDIZZONE: -- prescribe these as well, as do PAs.

CO-CHAIR THRAEN: Absolutely.

MEMBER ARDIZZONE: Thank you.

CO-CHAIR THRAEN: Albert?

MEMBER WU: I'd actually say it goes a little bit beyond doctor shopping. I think that there is a care coordination function here, and a lot of times no one's in charge. No one is aware of who else is on the team. And it does behoove anyone who's vaguely interested in population health to create a team and to establish who's on it and so forth.

So I think that it will push us a little in that direction, which is good.

CO-CHAIR THRAEN: All right. I'm going to call for the vote. We're talking about the evidence.

MS. QUINNONEZ: We are now voting on --

CO-CHAIR THRAEN: Oh Chris. I'm sorry, I missed Chris.

MEMBER COOK: I was supporting that same fact, that as a pharmacist, what happens is, you find out by the health plan, which helps to regulate in finding out whether you're having overlapping days' supply from another place, whether you're having multiple prescribers.

And so what this does is it takes it out of that opioid database that's so hard to deal with, that people can't get into, and with the health plans, you get that instant point of care that's going to allow you to see what else is going on instantly with those alerts. And so it's very helpful in that regard.

CO-CHAIR THRAEN: All right. I'm sorry. Go ahead. Call for the vote.

MS. QUINNONEZ: We are now voting on Measure 2950, Use of Opioids from Multiple Providers in Persons without Cancer. We're voting on the evidence. Option one, high, option two, moderate, option three, low, option four, insufficient. Yes, 2950.

CO-CHAIR THRAEN: All right. So we're on moderate. So go -- we have to re-vote. Re-vote. I guess, honest, Missy. Thank you.

MS. QUINNONEZ: Okay. We're re-voting on Measure 2950. The criteria has changed. Option number two is moderate. Option number three is low, and option number four, insufficient.

Option number two, moderate, option number three, low, and option number four, insufficient.

Okay. We have all votes, and voting is now closed. For the evidence of Measure 2950, 100 percent voted for moderate.

CO-CHAIR THRAEN: Performance gap.

MEMBER ARDIZZONE: I'm sorry. I think I'm the lead on this. So as with the measure before, they demonstrated a performance gap across the three different health plans that they looked at.

They also reported some disparities when they looked at the participants who were in the low-income subsidy. There was a really big difference in their usage rate per 1,000, as compared to people who do not get the LIS subsidy, indicating a performance gap.

CO-CHAIR THRAEN: Questions? All right. Call for the vote.

MS. QUINNONEZ: Voting is now open for performance gaps of Measure 2950. Option 1, high, option 2, moderate, option 3, low, option 4,

insufficient.

Looking for one more vote. Okay. All votes are in. Voting is now closed. For the performance gap of Measure 2950, 65 percent voted high, 35 percent voted moderate, 0 percent for low and 0 percent for insufficient.

CO-CHAIR THRAEN: Reliability.

MEMBER ARDIZZONE: So as with the other measure, I think there's going to be some discussion about the denominator. The numerator is different, though. Again, they're looking for any member with four or more unique pharmacy providers, and four or more unique prescribers. They're in the numerator.

In the denominator was the discussion that happened before. I do support that denominator, because I think it makes it more precise, so you have less noise. You have actually people who are at higher risk, instead of capturing all the people who get prescriptions, because you may be getting a high number of people that may not be actually fitting the criteria.

They do provide a list of opioid medications. It doesn't look like they're missing anything. Again, they're excluding patients who have cancer, and hospice.

Again, I'll make the statement again about patients who have cystic fibrosis, sickle cell or HIV, but I understand there were robust discussions among their committee members, and they did provide some reliability testing. Again, their results are pretty good.

I think they just, again, did them in a Medicaid population, and the mean reliability score was 0.93.

CO-CHAIR THRAEN: Questions? Yanling?

MEMBER YU: Oh, question is, how do you -- maybe you mentioned. How do you identify the providers that, you know, they shopped for, you know, getting mod. Do you have a plan to incorporate PMP, those types of data, at all?

DR. EISENBERG: The providers are identified by NPI number, presently. And is your

question, do we intend to include the information from prescription drug monitoring programs?

MEMBER YU: Yes.

DR. EISENBERG: PDMP?

MEMBER YU: Yes.

DR. EISENBERG: No. That wouldn't add anything to our measure, because our measures are all based upon claims, which we have captured to a very high percentage of reliability.

The PDMP information is frankly far inferior. You've heard some of the reasons for that today.

MEMBER YU: Yes. Okay. Thank you.

CO-CHAIR THRAEN: Other questions? All right. We'll vote.

MS. QUINNONEZ: Voting is now open for reliability of Measure 2950. 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 2950, 45 percent voted high, 55 percent voted moderate, 0

percent voted low and 0 percent voted insufficient.

CO-CHAIR THRAEN: Validity.

MEMBER ARDIZZONE: As discussed in the last measure, they did face validity only, which is okay, but only lets them get to a level of moderate.

They used their expert technical panel, which again, as we talked before, I cross-checked, and there's only one person who's on that technical panel who is a member of PQA, so I don't think there's any bias there. And they had 67 percent of their QMEP members who voted, on the face validity, who agreed.

In addition, they took all their 89 members to vote on whether to endorse the measure. And about 70 percent of them agreed that they should endorse the measure.

Threats to the validity, I think we talked about this. I'm sorry. I have nothing to say.

CO-CHAIR THRAEN: Questions? All right. We'll vote.

MEMBER ARDIZZONE: Voting is now open

for validity of Measure 2950. Option 1, moderate, option 2, low, and option 3, insufficient. Option 1, moderate, option 2, low, and option 3, insufficient.

All votes are in, and voting is now closed. For the validity of Measure 2950, 95 percent voted moderate, 0 percent low, and 5 percent insufficient.

CO-CHAIR THRAEN: Feasibility.

MEMBER ARDIZZONE: Feasibility seems easy to do. It's easily collected administrative claim data, and there are no concerns.

CO-CHAIR THRAEN: Questions? Vote.

MS. QUINNONEZ: Voting is now open for the feasibility of Measure 2950. Option 1, high, option 2, moderate, option 3, low, and option 4, insufficient.

Looking for -- all votes are in, and voting is now closed. For the feasibility of Measure 2950, 90 voted high, 10 percent voted moderate, 0 percent for low and 0 percent for insufficient.

CO-CHAIR THRAEN: Usability.

MEMBER ARDIZZONE: As discussed in the last measure, it's not currently publicly reported. However, it's part of a monitoring program for Medicare Part D, and CMS has announced plans to move this measure into a display measure for 2019, which would be publicly reported.

CO-CHAIR THRAEN: Questions? All right, we'll vote.

MS. QUINNONEZ: Voting is now open for usability and use of Measure 2950. Option number 1, high, option number 2, moderate, option number 3, low, and option number 4, insufficient information.

All votes are in, and voting is now closed. For usability in use of Measure 2950, 50 percent voted high, 45 percent voted moderate, 5 percent voted low, and 0 percent voted insufficient information.

CO-CHAIR THRAEN: All right. Suitability for endorsement.

MS. QUINNONEZ: Voting is now open for

the overall suitability for endorsement of Measure 2950. Option number 1 is yes. Option number 2 is no.

Option number 1 is yes, and option number 2 is no.

All votes are in, and voting is now closed. 100 percent voted yes for the overall suitability for endorsement of Measure Number 2950.

CO-CHAIR SEPTIMUS: Okay. We have a choice here. Our developers are very kind, and would be willing to come back for more torture in the morning. The next measure is a variation of the first two.

I don't know how long it would take us to get through that, but we do have, also, we have to ask for public comment as well, which, you know, usually goes fairly quickly.

So I'll leave it up to all of you whether or not you want to stick it out a little bit longer and try to get through the last measure, or whether or not you want to come back in the morning and start with this first thing.

Based on what we know about the first two, this is just a composite, really, of multiple providers at high doses, so it's a variant of the first two. So it really is the committee's choice.

CO-CHAIR THRAEN: All right, Number --

CO-CHAIR SEPTIMUS: Ms. Co-Chair, go for it.

CO-CHAIR THRAEN: -- 2951, and it's the Use of Opioids from Multiple Providers at High Dosage in Persons without Cancer. And Steve is the lead. And you want to say a couple?

MEMBER LAWLESS: The only thing I would say is that this combined measure is precisely mirroring the present Medicare over-utilization monitoring program. These are the patients that right now are being contacted by Medicare health plans.

Yes. I was going to say, but this is just an extension. This is the worst of the worst, in terms of what you're looking at. So yes, you're right. Your urine can be dropped off over there, if you want.

So anyway, the measure is a process measure, obviously. Inherently, it's looking at the worst that way. There is some conflicting evidence that you have, in terms of this, that we have four providers and this.

There are one or two articles you referenced, which actually said, in this particular measure, there may be less usage with these stopgaps. But that could have been just random chance.

So there'll be conflicting evidence that way. The bigger role, really, not addressed here, this is really more of a -- and I think you just answered it for me. We're finding people who are addicted, who are searching.

Because if you look at the rates that they have, you've listed here, in terms of what their mean need is and everything else, that almost mimics the rates, what I've seen published on how many people in the country are addicted.

So if this is looking more or less of a provider, more of as a screening of what is your

patient population, you need help, I think, and that's what the intent is.

CO-CHAIR THRAEN: Go ahead, Laura.

MEMBER ARDIZZONE: One quick question. Has there been any consideration to making sure -- the first two that we reviewed, you had some good data for a year or two, before combining them to, combining them together into another third measure.

DR. EISENBERG: I'd say two things. One is that there is evidence that all of these are independent risk factors, and that when they are put together, they are really identifying high risk patients, as we just heard.

And by the way, it's not just addiction. It's also redistribution of drugs, right. Some of this is just, you know, lawlessness.

The other thing is that the health plans want to identify the worst offenders. And this is a way to identify the worst offenders. We know that, based upon the over-utilization monitoring program, which is a retrospective drug utilization review program, we know that it can be effective.

So that's the sort of ground work that's been done for it. And naturally, we'll be collecting data over the next years, to define that better.

MEMBER ARDIZZONE: All right, just a follow-up. My question wasn't questioning the evidence for collecting the data. I meant the reliability and validity, feasibility, usability of these new measures, making sure that maybe for a year, they're actually capturing what you want.

They're easy. They're really precise, so that when you combine the two of them, they're the strongest that they could be. That's all I was asking.

CO-CHAIR THRAEN: She's referencing the past experience with composites, basically.

MEMBER LAWLESS: So along with that, in terms of the, does the measure capture it as -- your evidence you present talks about, obviously, the complications of narcotics, you're going to find, when in truth, really, you don't mention much.

But it really truly is about diversion.

It's about finding addictive behaviors. And I would focus a little bit more on that, because that also helps it, makes it justified.

And the only other correction I would make for you, I don't think this is a sign of -- especially with my last name, a sign of lawlessness. So I would actually, if we could take that out of the minutes, my family would appreciate that.

(Laughter.)

DR. EISENBERG: My apologies.

CO-CHAIR THRAEN: Okay. Kimberly?

MEMBER APPLEGATE: Yes, and your point's well taken about, you know, do we want to consider, also I had that comment, too, about waiting and getting the data right, and tweaking it.

And the other point is what you said. Are we being punitive in looking at this? Are we trying to be punitive? You know, I don't prescribe opiates, but I know a lot of people who do. And is the goal to be punitive to others, or in the name of quality improvement, are we trying to help patients?

And there are a lot of people out there that are suffering. And the goal, to me, is to find alternative, and push health care systems to help patients.

The VA system has failed patients many times over this issue, and has failed patients in providing enough providers in behavioral health, and they still haven't fixed it.

Over and over again, we see failure, and we're pushing this, these measures. And I think they're good measures. What I don't see happening is fixing the other half of the problem.

So I just caution everybody to say, okay, we're going to get the bad out, and it sounds punitive. And I want to remind us all that we want to help patients, and we want to help them get the help they need, not just reduce opioid use.

So I'm just asking us, that we're, we're going to cut out waste, and administrators love this. What we're not doing is getting the other half of the picture.

CO-CHAIR THRAEN: Thank you. Albert,

and then Missy.

MEMBER WU: I was going to say almost the same thing, and that is, you know, the measure might be the proportion of these people who you identify who get into a drug treatment program within six months. I mean, you know, honestly.

So I'm not completely clear on the goal of instituting this measure. We may find that, you know, the Hopkins program has a ton more opiate addicts than the one in Utah, but is that helpful?

DR. EISENBERG: I think it's helpful, because by identifying these patients, and by identifying all of their prescribers, because that's the information that's going to be generated, you, as a health plan, will be able to contact both your prescribers and your members, and begin a dialog that maybe hasn't happened.

CO-CHAIR THRAEN: That's what happens. Missy.

MEMBER DANFORTH: I think there is a typo in the measure sheet. It looks like, if you scan down past the evidence, it's the exact same as

the first measure. And it doesn't match what's later in it.

MS. BUTTERFIELD: I was going to actually point that out, because I noticed that, too. On the NQF face sheet, that is the case, if you look under, if you look under actual submission, it's different information. And if you look on the NQF face sheet under 2B5, that is the correct information. So I think that might have just been --

CO-CHAIR THRAEN: So it's wrong in one place, and right in --

DR. EISENBERG: Yeah, it's probably --

MS. BUTTERFIELD: It's wrong on the face sheet.

DR. EISENBERG: I understand. Yes.

MS. BUTTERFIELD: But it's correct in our submission. Yes. Thank you.

MS. PEZZULLO: And if I could just comment additionally on this third measure, the interest of the measure development workgroup in having this third measure was also in recognition

of the different levels of resources that health plans will focus.

And you know, I think, in agreeing that, having high patients using opioids at high dose is a concern, and also patients who are getting these prescriptions filled from multiple prescribers and multiple pharmacies is an issue. And when you combine both of those, it's, a serious concern.

And so for, you know, plans where they may have limited resources to dedicate to these efforts, this kind of brings this population, elevates this population so that they can dedicate their resources towards, this most at-risk population.

So, the primary intent of the measure development group was around the safety aspects. And of course, when we look at this, just by nature of focusing in these areas, you also end up addressing some of the diversion or misuse as well.

CO-CHAIR THRAEN: Lisa?

MEMBER MCGIFFERT: I'm not quite sure how to say this, but I do feel like this is about

providers enabling those patients, too. I mean, there are some responsibilities for providers, prescribers, to make sure their patients aren't already taking drugs from other prescribers. At the pharmacy level, there's a responsibility, especially if it's in a plan pharmacy.

It just seems to me that we have to get at the core to get at this problem. We have to get at the professionals that are enabling these, some of these patients to have ridiculous amounts of prescriptions.

And I'm not talking about somebody who has a gunshot wound, something like that. But I don't think that's what we're talking about here, and I just don't think it's really going to, it's going to capture the real problems, it seems to me.

CO-CHAIR THRAEN: Albert, do you have your -- no? Tracy?

MEMBER WANG: So at a health plan level, we have the data. And so there are ways to intervene. So, you know, so speaking for my own health plan, we have implemented a pharmacy home

program, whereby we identify these high risk members who are utilizing more than, you know, the necessary opiates, and also prescribers who we're also able to identify the different prescribers who have contributed to the over-prescription.

And we send a letter out to the providers, so that they can help, so letting them know that this is your member. They're using, you know, pharmacy scrip from, you know, XYZ places, can you do something to help reduce the overuse. So, you know, there are things that we can do to help them out.

CO-CHAIR THRAEN: Any other comments? Ed looks like he wants to say something.

CO-CHAIR SEPTIMUS: I'm just sitting here listening to this discussion. And, you know, it just so happens, in this week's New England Journal of Medicine, it talks about opiate treatment. Is there any doubt in your mind about the number of accidental deaths that occur, overdoses of opiates?

MEMBER APPLEGATE: Is there any doubt

in your mind about how many suicides there are in our VA vets, because they're not getting psychiatric care?

CO-CHAIR SEPTIMUS: I think that you raise a legitimate point, but I think there's a great opportunity, through these measures, to help.

And I think that there's a great opportunity for us to learn together to use medications appropriately. And yes, is to get them into the right care settings, to address their addiction. But I think --

MEMBER APPLEGATE: I'm here to help you.

CO-CHAIR SEPTIMUS: Thank you.

(Laughter.)

CO-CHAIR SEPTIMUS: See. I mean, but unless we identify these folks through some mechanism, then these folks will continue down the same path. So I'm just sitting here listening to this, saying this is a real major issue. And yes, these measures will not cure the problem, but it'll be an important first step to identifying who's at

risk. Yes. And --

CO-CHAIR THRAEN: So I want to also talk, just remind us that we're moving towards behavioral health integration.

CO-CHAIR SEPTIMUS: Right.

CO-CHAIR THRAEN: So the psychologists at the table, the social workers at the table, we're now being invited to join you in your health care delivery system. And, you know, we're at the very beginning of what that looks like.

But those are the resources that are coming to the table to try and help inform. You have to figure out how to identify them, and have to, you know, step out of the enabling role, but we're the ones that bring the interventions to the table.

All right. Martha, and then somebody else?

MEMBER COOK: Chris.

CO-CHAIR THRAEN: Chris.

MEMBER DEED: I just wanted to say that we just had an experience in Buffalo, which I could

talk about for hours, but I won't. But I almost think that the Buffalo experience should be appended to some of these measures.

We had the greatest prescriber of pain medication in the state in Buffalo. He was arrested. His practice was shut down. He had 10,000 patients. It resulted in suicides, break-ins into hospital pharmacies, local pharmacies. Our doors had to be locked at all times. They eventually got the Health Department to intervene.

The point is, they arrested the guy without giving any consideration to the 10,000 patients, granted, some of them addicts, unfortunately, some of them legitimate patients. And it's been an ongoing horror, an absolute nightmare for thousands and thousands of people and families.

That's not to say these measures shouldn't be put into place. You haven't heard word one out of me about that. But it is a really important public health consideration to consider

how you implement these things.

CO-CHAIR THRAEN: Chris.

MEMBER COOK: Sorry I'm breaking in line in front of Yanling.

MEMBER YU: Go ahead. Go ahead.

MEMBER COOK: We all know where CMS is going, and they've given us the road map. And the alternative pavement model's in the direction we're headed.

As we move down towards that capitated model and what's there, we have to get out of the silos that what we see is our traditional health care system is what it is, and that the social system is completely different.

As you start looking at a totality, all the stuff that we do and all the brainpower that's in this room and traditional health care makes up 20 percent of health, according to the World Health Organization.

So as we move towards that, to where we're looking at broader accountability, we're already starting to see those health systems

looking at those social determinants of health, and looking beyond what is just within their silo.

So I guess my optimistic, glass-half-full, is as we move there, we're going to see those things that are in -- become investments towards true outcomes of care, where we've ignored those in the past because they weren't within our silo.

So this is only providing further information for us to get to who has real issues. The next step then would be, how do we advocate on the patients' behalf to get them into those behavioral programs, into those things that actually assist, and the ones, whether it is pure diversion, or helping our criminal justice system, that is, helping to get that out of the way that our resources are used more efficiently and effectively. Sorry, and there's my soapbox.

CO-CHAIR THRAEN: Second. Yanling? And then Kendall.

MEMBER YU: Just a comments on the, how to bring everybody onboard, the physician's sides

and patient, families, on this, reduce the harm due to, you know, overuse of opioid.

You know, we just talk about a measure today about elderly using risky medication. And there's a credentialing building in the measure, for the physician, and for the facility, whatever it is.

So I was just wondering if, down the road, if you're looking at, you know, physician education for the whole population or for the physician, you could building in some type of a credentialing that might as be in there, as a motivation to really help change the behavior of the prescribers.

DR. EISENBERG: That's not something that my organization would do, but Secretary Burwell has an extensive outline of a plan that's been laid out, and physician education, and patient education are big parts of that plan.

MEMBER YU: What about the credentialing, physician credentialing?

DR. EISENBERG: Credentialing?

MEMBER YU: Yes.

DR. EISENBERG: I --

MEMBER YU: The entry into a --

DR. EISENBERG: Yes. I can't comment on credentialing. That would be up the boards of pharmacy for pharmacists and boards of --

CO-CHAIR THRAEN: And state licensing does some of that. Each state is different, so but the state license the use of the opioid controlled substances, and at least in the State of Utah, you're required to do the webinar type of thing, as part of the training, in terms of using -- and each year we're trying to increase and upgrade that training.

Randall?

MEMBER WEBB: Kendall.

CO-CHAIR THRAEN: Kendall. Sorry.

MEMBER WEBB: That's okay. So I just want to go on the record. I know it sounds like I am advocating for prescription opiates. I'm not. I am known by my residents as the narc Nazi. I am merely, as somebody who works in a very nasty urban

setting, trying to keep a hospital open that's providing care for a set of patients.

And I know there's lots of us all over the country, and I just don't want something like this -- what I'd like to see I something like this measure to create more what Iona was talking about, something proactive, something to give us something we can use, not the databases we have now, that are no good, that don't help us.

I live on the Georgia border, and I have friends who live in Pensacola. And they have four states to choose from in Pensacola. So what database do you look at? And how much time does it take to look at all four?

So I would love to see this measure or measures like this create a positive change, helped by the government, helped by CMS, to allow us to create something proactive.

As I'm ordering, I go to order, you know, Norco, because I don't order Percocet, and, you know, it tells me, oh wait, this guy had, you know, a Norco prescription three days ago. Great. Now

I know, and I can go back and say hey, you lied to me. You're done.

But without that, I think, this is exactly where we need to go. We need to go more towards a plan, until we have tools that the physicians and the practitioners can use to be able to make better decisions. We just don't have the tools right now.

MEMBER LAWLESS: Real quick, maybe a suggestion is changing the name of this. You see the passion in what everybody jumped on, as we're reading this and seeing it. When I first read it, it was more like, so here it is, the plan's going to go after the providers, and you've over-produced, over-prescribing.

But my suggestion would be as changing the name of it, to fit more what the conversation is. And I think you'll see that makes it a little bit more like what we're trying to do with this, rather than looking at providers.

DR. EISENBERG: Great. Thank you.

CO-CHAIR THRAEN: All right. We need

to vote, you guys, on the evidence.

MS. QUINNONEZ: We are now voting on Measure 2951, Use of Opioids from Multiple Providers and at High Dosage in Persons without Cancer. Option number 1 for evidence is high. Option number 2 is moderate.

CO-CHAIR THRAEN: It's the old one. It's the other one. Yes.

MS. QUINNONEZ: All right. Cancel that. Yes, it is. We're here now.

PARTICIPANT: We have to start all the way from this morning.

(Laughter.)

CO-CHAIR THRAEN: We're not going there.

MS. QUINNONEZ: We're ready to vote on the evidence for Measure 2951. Option number 2 is where we'll start, which is moderate. Option number 3, low, option number 4, insufficient. So option number 2 is moderate, option number 3, low, and option number 4, insufficient.

Okay. All right. All votes are in.

Voting is now closed. We have 0 percent for -- oh, obviously, 94 percent for moderate. We have 6 percent for low, and 0 percent for insufficient.

CO-CHAIR THRAEN: Performance gap.

MEMBER LAWLESS: In terms of adding, there's a little bit of a performance gap around the variation, in terms of signal-to-noise, if I read it correctly. I think the bigger performance gap we're talking about is resources for the patients and resources for the systems who take care of these patients.

So if this identifies that as a performance gap, it's a home run.

CO-CHAIR THRAEN: Okay. So we have to ignore what he just said. Any -- Missy?

MEMBER DANFORTH: Yes, just a question. So for the hospital and other provider-level measures, I think it's more clear of what to do when there are disparities identified. This measure actually has a huge disparity that was identified with one plan that looks at low income folks.

And so when you see a disparity that's

like so significant like that, how does that go into evaluating the measure? I mean, we're looking at a falls measure, and there's a disparity for older women in particular, right. We talk about adjustments.

We're looking at readmission measures and there's disparities, right. We talk about all kinds of facility-level adjustments. This is a health plan level measure where a huge significant disparity was identified. So how does that go into, you know, our processing of the measure?

Specifically a health-plan level measure, right, because that's what I don't understand.

MR. LYZENGA: I mean, I would think that that would speak to a larger opportunity for improvement. But again, this is one of those ones that's open to interpretation, this opportunity for improvement category.

DR. EISENBERG: Our approach to this is through stratified reporting. We recognize that there are huge disparities amongst the different

lines of business. We think that Medicare/Medicaid/commercial needs to be recorded separately.

It might more need to be done beyond that. We'll learn.

MEMBER DANFORTH: Real quick on that, though. So if by virtue of having certain populations of people including in your plan, your performance on this measure is worse, I would think that we would, instead of stratifying the reporting, we'd want to do something to make sure that those plans were doing something extra to acknowledge that they were having this problem.

I'm just trying to draw parallels with other types of measures that we do. So stratification would make sure that they're compared to each other, right, in a fair way, but if we're really going to sort of drive change and improvement, I would just think that we would want to do something else besides stratification when we see that kind of disparity.

So for example, when we're looking at

readmission measures in hospitals, if we saw that kind of difference between hospitals in different communities, we'd form an NQF committee to look at socioeconomic, right, to adjust the measure and do all these extra kinds of things.

So I'm just thinking like, I just feel like there's sort of a, maybe even a moral obligation to look a little bit closer when we're identifying that kind of significant disparity for a certain population of people, that really the health plan has the power to identify to the name and address level. I guess that's where I'm going with this.

DR. EISENBERG: We agree. We've got work to do.

CO-CHAIR THRAEN: Vote on performance gap.

MS. QUINNONEZ: We are now voting. Voting is open for performance gaps of Measure 2951. Option number 1 is high, option number 2, moderate, option number 3 low, and option number 4, insufficient.

Okay. Voting is now closed. 63 percent voted high, 38 percent voted moderate, 0 percent for low, and 0 percent for insufficient.

CO-CHAIR THRAEN: Reliability.

MEMBER LAWLESS: We're on reliability. The reliability testing, actually, was 0.92, which is good, I mean, is very strong. But it was only performed within the same group, the same measure, but it's pretty straightforward, claim data, very easy to reproduce and stuff. So I think the reliability, it shows the reliability as high.

If one asked more than just what -- it doesn't go into the appropriateness of things like documentation, state characteristics, use of medical marijuana in certain states, and how that would impact this or not.

And it also, which you have brought up, the reliability of multiple plans and multiple locations in multiple states. So it's a limitation that would very well -- otherwise, it's good.

CO-CHAIR THRAEN: Any questions? All right, we'll vote.

MS. QUINNONEZ: Okay. Voting is now open for the reliability of Measure 2951. Option 1, high, option 2, moderate, option 3, low, and option 4, insufficient.

All votes are in. Voting is now closed. For reliability of Measure 2951, we have 69 percent voted high, 31 percent voted moderate, 0 percent for low, and 0 percent for insufficient.

CO-CHAIR THRAEN: Validity.

MEMBER LAWLESS: I have nothing new to add to what we've already talked about, you know, to validity.

CO-CHAIR THRAEN: Any questions? Vote.

MS. QUINNONEZ: We're -- voting is now open for the validity of Measure 2951. Option 1, moderate, option 2, low, option 3, insufficient.

Option 1, moderate, option 2, low, and option 3, insufficient.

All votes are in, and voting is now closed. For validity of Measure 2951, 89 percent voted moderate, 11 percent voted low, and 0 percent

insufficient.

CO-CHAIR THRAEN: Feasibility.

MEMBER LAWLESS: Again, as we talked about, nothing new to add from what we've already talked about, and hammer this.

MS. QUINNONEZ: Voting is now open for feasibility of Measure 2951. Option 1, high, option 2, moderate, option 3, low, and option 4, insufficient.

CO-CHAIR SEPTIMUS: You see, you've worn him down.

(Laughter.)

MS. QUINNONEZ: Voting is now closed. For feasibility of Measure 2951, 88 percent voted high, 12 percent voted moderate, 0 percent low, and 0 percent insufficient.

CO-CHAIR THRAEN: Usability.

MEMBER LAWLESS: Even though we've also talked this one down as far as we possibly can, but I think the usability as it's presented in here, in terms of use, is different from what we've been talking about.

And so I think the emphasis of usability -- somehow I have to tell you, I have a disconnect here, because we've been talking about the passion, and what we're talking about how we really could use this. In reading all the details of the document, that's not what comes forth when you read the measure.

So if I'm going by the measure of why complications over use, that's what's in this document, not the idea of behavioral health, watching for people who are addicted.

CO-CHAIR SEPTIMUS: You want to comment on that?

DR. EISENBERG: Yes. I think the usability of this measure is really quite high, because we know it's going to be identifying patients and their prescribers that are, together, leading to high doses of medications for prolonged periods of time from multiple prescribers. It gives multiple avenues for intervention.

And we know that the data is already being collected for an opiate over-utilization

monitoring program. So to us, it works.

MEMBER LAWLESS: And I know those of you, and I'll stop after this, if we -- other measures, when we've actually looked at what the measure has been presenting, and you look at the measure presenting, the outcome you're looking at is the thing you're looking at where your validity is, overdoses and physiologic complications of the opioids.

There are the other aspects of this, too, which is the identification of systems needs and stuff like that. So I just, again, it's very usable, very, very usable. But I think, in terms of this, I have a disconnect.

CO-CHAIR THRAEN: Well, I'd like to make the argument that these measures will help formulate policy. Because I think what we're talking about is, what do we do once we've identified the problem.

And because the problem's been pretty much shown up in the criminal justice system or the ED system and not really at the plan level, when the

plan starts to get a handle on what percentage of their population is causing these kinds of issues, then that becomes a positive decision as to what they're going to do, if they're going to either kick them off the plan, or they're going to intervene, or they're going to do something, you know, to address that once they've identified it

So I would make that argument that what we've been talking about with behavioral health is really at the policy level decisions that this measure could inform.

MEMBER YU: I'm serving on the State Medical Board in Washington. We have reviewed how multiple patients got killed over the years, by a physician who prescribed pain medication really irresponsibly.

But unfortunately, those tragic events only be known after multiple patients got harmed and dead. So I just wonder, is there any chance, down the pipe, that there will be involved with the health care plan, would inform whoever, medical board or whatever, inform those very risky

prescribing pattern that involve multiple patient harm.

The agency work together to really protect the public safety.

DR. EISENBERG: Well I wish I could tell you that that was going to happen easily. But my experience, working as a medical director at health plans for many years, is that our contacts in the criminal justice system have not been so welcoming of our information.

They often times will listen to us, and then we get no response from them. So we generally don't know what is done with the information that we've provided to the criminal justice system, or to the boards of pharmacy or medicine.

So the information will be there. We'll have more information. We'll have better information than we've ever had before. What other organizations, especially government organizations, will do with it, we don't know.

MEMBER YU: But your data would be public, right?

DR. EISENBERG: Yes.

MEMBER YU: Will not be public?

DR. EISENBERG: No, no. It'll be -- depends on the system that it's in. In the Medicare system, it will be public as soon as it's a display measure, which is, the data for that is starting to be collected in 2017, for 2019 publication.

MEMBER YU: Will that be public to the state agency?

DR. EISENBERG: Yes. It'll be public to everyone, yes.

MEMBER YU: Okay. Okay, thank you.

CO-CHAIR THRAEN: Lisa?

MEMBER MCGIFFERT: My view of this measure is, it is a process measure, and as all process measures, you're trying to effect a change in behavior, or you're trying to make something good happen, or make something bad not happen, like an infection.

So it seems to me that it does have a capacity to get us at least to a certain point. It doesn't have the capacity to fix the whole system,

but as a process measure for this specific behavior, it seems appropriate.

CO-CHAIR THRAEN: So I think that one of the things you start to see is that if health plans identify a prescriber that's outside of the boundaries, and in their process of coaching the prescriber, the prescriber does not come within the boundaries, the prescriber will be let go, and they'll show up in another health plan, or across the river in the other state.

So one of the challenges has been, the DOPL, the Division of Public -- of Professional Licensing, the state government entity, has to figure out how they're going to work with health plans, so that one, there's an intervention that's done with the provider, if the -- the prescriber, excuse me, if the prescriber is themselves a drug addict, which often is the case, and/or is it just a criminal behavior, and distinguishing between that.

But those systems have not played together, historically.

CO-CHAIR SEPTIMUS: I don't think NQF can solve all --

CO-CHAIR THRAEN: No.

CO-CHAIR SEPTIMUS: -- the ills --

CO-CHAIR THRAEN: No, but we'll have data. Let's vote on usability, please.

MS. QUINNONEZ: Voting is now open for usability and use of Measure 2951. Option number 1 is high, option number 2, moderate, option number 3, low, and option number 4, insufficient information.

All votes are in and voting is now closed. For the usability and use of Measure 2951, 53 percent voted high, 47 percent voted moderate, 0 percent for low, and 0 percent for insufficient information.

CO-CHAIR THRAEN: Okay. I'm thinking this is the last vote of the night. Suitability for endorsement.

MS. QUINNONEZ: Voting is now open for overall suitability for endorsement of Measure 2951. Option 1 is yes. Option 2 is no.

CO-CHAIR SEPTIMUS: There actually is one more vote before we leave, and that's who wants red and who wants white.

(Laughter.)

MS. QUINNONEZ: All votes are in for the use, for the overall suitability for endorsement of Measure 2951. One hundred percent voted yes.

CO-CHAIR THRAEN: Thank you all.

CO-CHAIR SEPTIMUS: No, no.

CO-CHAIR THRAEN: Thank you.

CO-CHAIR SEPTIMUS: Wait a minute. We have public comments to the committee.

CO-CHAIR THRAEN: Oh, I'm sorry. Public comment. Anybody on the phone or in the audience wishes to comment?

OPERATOR: In order to make a public comment, press star, and then 1. There are no public comments.

CO-CHAIR THRAEN: We have one in the room. Hold on.

MR. CONYERS: Good evening, at this point, everyone. I certainly understand that I am

standing in the way of your wine. But I'm Del Conyers. I'm Vice President of Quality and Compliance at the National PACE Association. I'm also a NQF alum, so bear with me. Don't be too hard on me.

First just say, I appreciate the difficulty that you all had in understanding the PACE model. I know many of you sort of struggled with understanding the dynamics of the IDT team, the nuances of the PACE populations, so I certainly appreciate that. I found myself in the same position that you were in this morning when I took the role.

I just wanted to point out that again, that the IDT is an integral part of the PACE model, and that it has 11 disciplines represented in terms of those who provide care to the frail elders.

In addition to that, the patient and care giver are really paramount in influencing care planning as well as having an impact on the outcomes that they are often faced with.

So I just wanted to make sure that you

understand the role of patient autonomy and self-determination in influencing outcomes, and in the light of that, consider, you know, despite all the preventive measures that take place, a lot of the locus of control falls on participants in influencing outcomes.

So I'd just like for this group to consider that, moving forward. With regard to assessment, assessment happens frequently and on the continuum of care, every six months, when they're changing status. So I just want the group to understand that, as well, going forward. And I'll proceed quickly.

Just with the gaps of care, I notice that the concerns raised for the pressure also measure, I think, also apply to the falls. When we talked about, you know, there's no evidence to demonstrate quality for PACE programs specifically, I think that applies across measures.

But it felt like that was not deemed or viewed in the same light when we got to falls. And I just want to point out that I think that those

concerns should be relevant to falls as well.

With regard to reliability, I just think that while signal-to-noise, I know we talk a lot about that and got some clarification, I think while real differences are demonstrated between the sites, the fact that there was no statistical test to assess whether the performance rates were statistically significant, that's also something that we should consider as well.

When we look at the PACE programs, there's a high degree of variation in the patient population. I think there are differences in outcomes, related to the maturity of the organization, the frailty, risk assessment.

So I think that because the measures don't discern good and bad care, because that wasn't done, because the sample size was so small, should also be considered.

Lastly, with regard to usability, I know that CMS say what it wants to fight, is part of the process. I certainly acknowledge that. But I'd be remiss if I said that, you know, given the

implications of what endorsement will be on the walls of NQF, I think, needs to be concerned.

I'm concerned that PACE is often compared to institutionalized settings of care. As someone pointed out, it's really not an apples to apples comparison, but we often find ourselves in that position, often compared to long-term care settings, nursing homes. But they're quite different.

So while we are okay, to some extent, with comparison of PACE programs, internally, I would just caution the implications of this measure being used to compare to other institutionalized populations.

So I'm off my soapbox. I hope you all consider that, moving forward, and I appreciate your time. Thank you.

(Whereupon, the above-entitled matter went off the record at 6:17 p.m.)