NATIONAL QUALITY FORUM
+ + + + +
NEUROLOGY PHASE II
STEERING COMMITTEE
+ + + + +
WEDNESDAY
OCTOBER 3, 2012

+ + + + +

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., David Knowlton and David Tirschwell, Co-Chairs, presiding.

PRESENT:

DAVID KNOWLTON, MA, New Jersey Health Care Quality Institute
DAVID TIRSCHWELL, MD, MSc, University of Washington Department of Neurology
A.M. BARRETT, MD, Kessler Foundation

WILLIAM BARSAN, MD, University of Michigan Health System
JOCELYN BAUTISTA, MD, Cleveland Clinic
RAMON BAUTISTA, MD, MBA, University of Florida, Jacksonville
GWENDOLYN BUHR, MD, American Medical Directors Association

GAIL COONEY, MD, FAAHPM, Hospice of Palm Beach County
JOHN DUDA, MD, Veterans Health Administration
JORDAN EISENSTOCK, MD, CPE, UMass Memorial Health Care
SAM FAZIO, PhD, Alzheimer's Association

RISHA GIDWANI, DrPH, Stanford University Medical Center
DAVID HACKNEY, MD, Beth Israel Deaconess

Medical Center
MICHAEL KAPLITT, MD, PhD, Weill Cornell Medical College
DANIEL LABOVITZ, MD, MS, Montefiore Medical Center
THERESE RICHMOND, PhD, CRNP, FAAN, University of Pennsylvania School of Nursing
JACK SCARIANO, JR., MD, FAAN, private practitioner
PETER SCHMIDT, PhD, National Parkinson Foundation
RAJ SHETH, MD, Nemours Foundation
JOLYNN SUKO, MPH, Virginia Mason Medical Center
JANE SULLIVAN, PT, DHS, MS, Northwestern University Feinberg School of Medicine
FREDRIK TOLIN, MD, MBA, FACS, Humana, Inc.
MARY VAN DE KAMP, CCC-SLP, RehabCare
SALINA WADDY, MD, National Institutes of Health

NQF STAFF:
HEIDI BOSSLEY, MSN, MBA
HELEN BURSTIN, MD, MPH

ANN HAMMERSMITH, JD
KAREN JOHNSON, MS
SUZANNE THEBERGE, MPH
JESSICA WEBER

ALSO PRESENT:

GREGORY BARKLEY, American Academy of Neurology*

CHRISTOPHER BEVER, American Academy of Neurology

GINA GJORVAD, American Academy of Neurology

JULIE KUHLE, Pharmacy Quality Alliance

DAVID NAU, Pharmacy Quality Alliance*

REBECCA SWAIN-ENG, American Academy of Neurology

CHRISTIE TEIGLAND, Inovalon, Inc.

JACQUELINE VANCE, American Medical Directors Association

*Participating by teleconference

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

C-O-N-T-E-N-T-S (CONTINUED)

C-O-N-T-E-N-T-S (CONTINUED)

C-O-N-T-E-N-T-S (CONTINUED)

C-O-N-T-E-N-T-S (CONTINUED)

C-O-N-T-E-N-T-S (CONTINUED)

C-O-N-T-E-N-T-S (CONTINUED)

C-O-N-T-E-N-T-S (CONTINUED)

C-O-N-T-E-N-T-S (CONTINUED)

C-O-N-T-E-N-T-S (CONTINUED)

C-O-N-T-E-N-T-S (CONTINUED)

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

9:01 a.m.

MS. JOHNSON: Well, good morning, everybody, and welcome to NQF's Neurology Endorsement Maintenance Project, Phase II. We really appreciate all of you guys coming out and joining us today on this soupy, yet not so hot, day, as the last time when you were here in June.

What we are going to do today, I just want to make sure everybody is aware of the project team. I am Karen. I am the Senior Director on the project. Down to my right is Suzanne, and Jessica is roaming around the room. So, that is Jessica. And then, here on my left is Helen Burstin. She is the Director of our unit, the Performance Measures Unit. And then, also here to my right are my esteemed Co-Chairs for the project, David Tirschwell and Dave Knowlton.

So, thank you, guys, for joining us.

As we did last time, we are going to start off the morning with introductions, welcomes. Ann, our General Counsel, is going to tell us what we need to do our introductions around the table.

MS. HAMMERSMITH: Thanks, Karen.

I am Ann Hammersmith. I am NQF's General Counsel.

I think most of you were at the last meeting, so you are familiar with this portion of the meeting. What we are going to do is go around the table once again, have you introduce yourselves, tell us who you are with, and make any disclosure that you wish to make. Just because you make a disclosure does not mean you have a conflict of interest. It is simply a disclosure.

Before we start, I want to remind you of a few things. There is no need to recount your CV. Please don't because we will be here all day and you will never talk about the measures.

(Laughter.)

What we are particularly interested in you disclosing is anything that is relevant to the topics that will be discussed today and tomorrow in the meeting. In particular, we are interested in consulting, speaking engagements, grant monies, research monies, if they are relevant to what is before the Committee at this meeting.

I also want to remind you that conflict of interest and disclosure is not simply financial. Many times Committee members will say, "I have no financial conflict of interest." A financial conflict of interest is part of the scenario here. But because of the kind of work that all of us do, you can also have a conflict or something that should be disclosed for an activity where you are volunteer, such as serving on a committee if it is relevant to what is before us these two days.

And finally, I want to remind you that you serve as an individual. You are not here as a representative of your employer or of anyone who may have nominated you. Occasionally, Committee members will say, in good faith, "I'm" So-and-So, "and I am here representing the American Society of" fill in the blank. And actually, you are not. You are here because you are experts. So, you serve as individuals.

So, with that, I will start with the Chairs.

CO-CHAIR KNOWLTON: I am Dave Knowlton. I am the Chief Executive Officer of the New Jersey Health Care Quality Institute, and I have no conflicts.

CO-CHAIR TIRSCHWELL: Good morning, everyone. Welcome back.

I am David Tirschwell. I am a stroke neurologist. I work at the University of Washington in Harborview Medical Center in Seattle, Washington. I do not have any

relevant conflicts.

MEMBER RICHMOND: I am Terry Richmond. I am a professor at the School of Nursing at the University of Pennsylvania.

Since our last meeting, I received funding from NIH and yesterday from National Science Foundation. I don't think there is any conflict. One of my studies does look at depression on psychological consequences, but it is all related to injury and not directly related to these measures.

MEMBER SUKO: I am Jolynn Suko from Virginia Mason Medical Center, accountable for neurosciences there. I have no conflicts of interest.

MEMBER LABOVITZ: I am Daniel Labovitz from Montefiore Medical Center in Bronx. I am a stroke neurologist, and have nothing to disclose.

MEMBER R. BAUTISTA: Ramon Bautista, University of Florida. Nothing to disclose.

MEMBER J. BAUTISTA: Jocelyn Bautista. I am an epilepsy neurologist at the Cleveland Clinic. I have participated with the American Academy of Neurology in writing evidence-based guidelines for epilepsy, nothing directly related, though, to the measures today.

MEMBER BARSAN: Bill Barsan. I am in emergency medicine at the University of Michigan. I have NIH funding to run the Neurological Emergency Treatment Trials Network, which does do clinical trials in seizures and other neurologic emergencies.

MEMBER DUDA: I am John Duda. I a movement disorder neurologist from the Philadelphia VA Medical Center in the University of Pennsylvania. I have research support from the VA and NIH and Michael J. Fox Foundation, which I don't think is relevant to today's topics.

I serve on the Scientific Advisory Board for the Lewy Body Dementia Association,

which may be relevant for the one cognitive issue. And I do with the national VA formulary leaders to guide use of the formulary in the VA, but I don't think that is necessarily relevant, either.

MEMBER VAN DE KAMP: Mary Van de Kamp. I am Senior Vice President of Clinical Operations for Kindred and RehabCare, and I have nothing to disclose.

MEMBER SULLIVAN: I am Jane Sullivan. I am a physical therapist. I teach in the Feinberg School of Medicine at Northwestern University in Chicago.

I have funding from NIDRR and the Department of Education and from industry, but it is related to stroke.

MEMBER BUHR: My name is Gwen Buhr. I am a geriatrician at Duke University, and I have nothing to disclose.

MEMBER TOLIN: Fred Tolin, Vice President at Humana. Nothing to disclose.

MEMBER KAPLITT: I am Mike

Kaplitt. I am a stereotactic and functional neurosurgeon at Well Cornell Medical College in New York, and I have nothing to disclose.

MEMBER SCARIANO: Jack Scariano. I am a practicing neurologist, and I will be talking about sleep studies and, also, patients who have that. And I don't read sleep studies and I don't treat sleep studies. I mean, I don't treat sleep patients.

MEMBER BARRETT: I am A.M. Barrett from the Kessler Foundation, where I direct the stroke rehabilitation research. I have funding from the Kessler Foundation, from NIH, and from NIDRR, and the Wallerstein Foundation for Geriatric Improvement.

I am a member of the American Academy of Behavioral Neurology Section, and within that, of the Clinical Practice Work Group that discusses consensus recommendations for behavioral neurology activities.

MEMBER EISENSTOCK: I am Jordan Eisenstock. I am a neurologist at UMass

Medical Center in western Massachusetts. I am also a Board-certified psychiatrist. I don't have anything to disclose, no conflicts.

MEMBER FAZIO: I am Sam Fazio. I am a developmental psychologist. I am from the National Office of the Alzheimer's Association in Chicago, and I have nothing to disclose.

MEMBER COONEY: I am Gail Cooney. I am Board-certified in neurology and hospice and palliative medicine, but practice exclusively in the field of hospice and palliative medicine. I have nothing to disclose.

MEMBER GIDWANI: I am Risha Gidwani from Stanford University Medical Center. I have nothing to disclose.

MEMBER SHETH: Raj Sheth at Nemours Mayo Clinic, Jacksonville, Florida, epileptologist. Nothing to disclose.

MEMBER SCHMIDT: I am Peter Schmidt from the National Parkinson

Foundation. I have nothing to disclose, but I would like to comment.

The National Parkinson Foundation is listed as a cosponsor on the Parkinson measures. That came as a surprise to us when we saw the submission. So, we are not involved in that.

MS. HAMMERSMITH: Okay. Thank you for those disclosures.

Do any of you have any questions of me or anything you would like to discuss with each other based on the disclosures this morning?

(No response.)

Okay. Thank you. Have a good meeting.

MS. JOHNSON: Thank you, Ann.

Can we go ahead and bring up this morning's slides?

I wanted to start out the morning with just a very brief overview, some housekeeping details. By now, you guys have

all figured out how to work your microphones. But, just as a reminder, once you have finished speaking, please turn your microphone off, so that it will work for the next person.

To signal your desire to speak, if you would raise your name tag and set it vertical, that way, our Chairs will know that, just like you all just did -- thank you.

(Laughter.)

So that we know that you would like to speak. We would appreciate that.

You should have been given clickers when you came in. So, hopefully, everybody has a clicker. You will use the clickers to register your votes as we go through the day.

Next slide, please.

Most of you I think remember how to do this, but, basically, you will use the keypad to register either a 1 for yes or a 2 for no in the appropriate set criteria, and then 1, 2, 3, or 4 for high, moderate, low, or

insufficient.

Who is running the voting? Is that Suzanne? Okay.

Suzanne's computer is the computer that has the receiver in there. So, point your clicker at Suzanne when you get ready to vote.

You will have 60 seconds to vote. If you are not sure that your clicker activated, just keep clicking your selection. It will not double-count your vote.

Just a quick overview. I am sure you guys already know this. You could probably recite this in your sleep. But we will be looking at 22 measures in Phase II, twelve on dementia, three on epilepsy, six on Parkinson's, and then one-off, stenosis measurement in carotid imaging studies.

Most are new measures. As a matter of fact, only the carotid imaging measure is an already-endorsed measure. So, everything else is new to us.

And you will have also noticed that most of the measures, 18 of them actually, have not yet been tested for reliability or validity. We have communicated to you a couple of different times on why we did accept those kinds of measures, that basically they are fairly non-complex measures. Let me think. They hit a gap area. So, in other words, we don't already have measures in the NQF portfolio that address the focus of the measures.

And they are also time-sensitive in a particular way. So, in the case of the measures for epilepsy, several of the dementia measures, and the Parkinson's measures, those will be used in the 2012 PQRS program. So, we consider that as a time-sensitive -- I don't know what the word is, but we consider it time-sensitive. And therefore, we did want to look at these measures.

As we go through and we talk about reliability and validity, it will be different

because, since they have been tested, you will not be thinking about how it was tested and was it at the measure score level or the data element level, all that kind of stuff, but you will still have to think about the specifications, particularly the precision of the specifications and, also, how those line up with the evidence. So, you will see that as we go through the day.

Next slide, please.

You have some tools to help you throughout the day. First of all is your meeting agenda. I believe that has been passed out to you.

We are planning to go in order of the agenda. As you know, sometimes things have to get moved around, but, in general, we are planning on sticking with the agenda.

We also have provided what we call our summary document. That document contains brief descriptions of the measures, comments from your preliminary evaluations, and then,

finally, the Work Group summaries that we came up with after participating and listening to your talks on the Work Group calls.

You also have our quick guide. The quick guide is a little four-pager that reminds you of all the different criteria, subcriteria, and the rating scales that you will use.

And then, of course, you have measure submission materials. You probably haven't printed those off, but they are available on the SharePoint site or perhaps you have already downloaded them to your computer.

Finally, there is one set of comparison tables for related measures, and we don't even need to talk about that until tomorrow.

Next slide, please.

Our process today is going to be pretty much the same as it was the last time around. We will discuss each subcriteria and

then vote. So, basically, we will talk about impact and then vote on impact, and then talk about evidence and then vote on evidence, et cetera.

You will notice that evidence we have numbered as subcriterion 1þ. We have switched those around, so we will talk about evidence first and then talk about opportunity for improvement. Part of the reasoning there is often measures have difficulty at the evidence subcriteria. So, if something is going to die, for lack of a better word, at evidence, we won't take the time to talk first about opportunity for improvement. So, it is just a time management strategy.

If a measure fails a "must-pass" criterion, we will stop. Okay? So, we won't go on to the other subcriteria. That is a little different than what we ask you to do in the Work Group calls because we wanted you to think about all of the criteria for the measure. But we will not be doing that here

in person today. So, if something dies on impact, we won't talk about any of the other criteria.

For our first measure, as necessary, we will review the evaluation criteria. So, I may jump in a little bit more on the first measure as you talk through it, just to remind you of what the rating scales look like or give you pointers about how to consider and evaluate the measure. I don't think I will need to do that very much throughout the rest of the day.

We will have roughly about 15 minutes per measure. As you know, generally, how it works is we are a little bit slower with the first measures, and then we kind of speed up throughout the day. But, on average, we are going to be looking at 15 minutes per measure.

Next slide, please.

Most of you have been assigned a role of lead discussant for measures. You

have pretty much had a chance to do this at least once now on the Work Group calls. But, basically, just a reminder, we want you to lead the discuss against the criteria. So, how did the measure stack up? We want you to summarize your thoughts and the thoughts of the Work Group and the discussion in the Work Group, particularly on how well the measure meets or does not meet the criteria. Okay?

We really hope that everybody feels free and comfortable to participate in the discussion of all the measures. Even if it is not your thing, please definitely chime in. And a reminder that the entire Committee will be voting on the measures and whether or not the measures meet the criteria.

Next slide, please.

So, let me stop there and see if there are any questions about process, housekeeping, et cetera.

(No response.)

Okay. Go to the next slide,

please.

I wanted to give just a very, very quick overview of the criteria. I know, again, you guys are old hands at this; you probably don't need this, but just in case.

Next slide, please.

Just a reminder that we have four main criteria for you to look at and evaluate measures against: importance to measure and report, scientific acceptability, usability, and feasibility. The first two are what we call "must-pass" criteria. So, again, a measure must pass importance before we go on to discuss scientific acceptability, and et cetera.

Next slide.

Under importance, we have three subcriteria: high impact, evidence, and performance gap or opportunity for improvement. And again, all three of these are "must-pass". So, we will, again, talk about these in order. But each of the three

subcriteria under this main criteria of importance to measure and report must pass.

Okay. Next slide.

Just a reminder in terms of thinking about evidence for the measure focus. In general, NQF does have a preference for certain measures, in particular, outcome measures. So, that is what we would love to see. But, of course, that is not always easy to do, and we don't always have a lot of outcome measures. As a matter of fact, in this phase we have no outcome measures for you to consider.

But, then, in order of decreasing importance or preference, we would also love to see measures that are intermediate outcomes -- often, those are kind of clinical-type outcomes -- or process or structure measures. Within those, we prefer those that are most closely linked to outcomes.

So, again, evidence is a very big deal here for us. I think possibly what you

will have seen as you look through the measures, often, if evidence is lacking or may seem to be lacking, it could be because the proximity to an outcome.

So, next slide.

This, again, is our little schematic just showing you that there are lots of different types of process measures. Again, we still have our hierarchy with preference for outcome measures and intermediate outcome measures. But, if we have process measures, the ones that we prefer are the ones that actually look at provision of intervention, and then, going backwards, actually choosing or planning interventions, identifying or diagnoses, and then assessing is even less proximal to the actual health outcome. So, this is just a reminder of that preference.

Next slide, please.

There is a difference between a low rating versus a rating of insufficient

evidence. So, low rating means that the evidence is there, and it didn't really demonstrate that the criteria has been met.

Insufficient evidence could be either that the evidence is there and presented, but still didn't answer the question, or perhaps the evidence is there, but it just didn't make it to the submission form. So, there's a couple of different ways that you could have insufficient evidence. In both cases, either low or insufficient evidence, a measure would not pass, but, again, it is for different reasons.

Next slide.

In terms of evaluating measures for the evidence subcriterion, again, we don't have any health outcomes. So, for all of the measures that you will be looking at today and tomorrow, we ask the developers to provide explicit and transparent information on the quantity, quality, and consistency of the body of evidence. So, again, body of evidence is

the entire body, not just selected articles.

What we hope that they are able to do, because it makes their life easier really, is if they can find evidence that has already been graded and collected, so that they can just report to you the summaries from those already-digested, if you will, reviews. We prefer grade or USP -- I can't even say the letters -- U.S. Preventative Services Task Force system. If they use a different system, we do ask that they describe what the different grades mean. I think the developers this time around did a great job on that part.

There are separate rating scales for quality, quantity, and consistency. I guess probably the final thing from this slide is just to remind you that expert opinion is not what we consider evidence. Okay? So, we are looking for empirical evidence.

Next slide.

In thinking about the opportunity for improvement, so that is subcriteria 1(b),

that we will do third under importance to measure and report, one of the things that we ask the developers to do, if they can, is to provide some information about disparities. Often, they are not able to do that. But today I am going to specifically ask you, as a Committee, if you have any information about whether a measure might be what we would call disparity-sensitive. This is to support kind of an ongoing process that we are getting ready to implement.

So, basically, I will just be asking you if you know whether or not this measure is possibly disparity-sensitive. And then, if so, do you have any sources that you could point us to for us to go and understand that literature? And you may or may not.

Next slide, please.

Just a quick reminder. We have the generic rating scale, high, moderate, low, and insufficient. Those will be used for subcriteria 1(a), 1(b), and for usability and

feasibility.

Next slide.

Evidence subcriteria, there are different rating scales for quantity.

Next slide. Quality.

And next slide. Consistency.

And then, go to the next slide. This is just the decision logic, which is pointing out that pretty much you need a moderate or high on all three of those in order to pass the evidence criteria.

Okay. Next slide.

I don't think we had to talk about the potential exception in Phase I, but we do have a couple of potential exceptions to the evidence subcriterion. Actually, we did talk about the first one in terms of health outcomes.

If you recall, for outcome measures, we didn't ask that developers tell us about quantity, quality, and consistency of evidence. We just asked about a rationale for

their outcome measure.

But, for other types of measures, non-outcome measures, we do have room for a potential exception to the evidence subcriterion. Basically, if there is no empirical evidence at all, but there is expert opinion that has been systematically assessed, and you also feel that the benefits would outweigh the potential harms, then you could consider invoking this exception to the evidence subcriteria. Okay?

So, if that becomes an issue, then somebody around the table would probably want to say something about "I think we should discuss invoking the evidence exception." If there is kind of general consensus around the table, then I think we vote on actually applying that exception. Okay? Anybody have any questions on that piece?

(No response.)

Okay. Next slide.

This is also a reminder. Our

Consensus Standards Approval Committee a while back did come out with some guidance for measure construction. I wanted to share just a couple of things that they found. These are the folks that see all of the measures from all the projects. Because all the Steering Committees do their thing, and then we take things to the next level, which is CSAC.

So, the CSAC folks see everything and they create some guidance for us. Basically, some of the guidance that they put out for developers to reflect the evidence criterion is to avoid measures that can be met primarily through documentation. That is one of the things that they suggested doing.

A lot of the times we use the term "checkbox" measures. So, those generally aren't the kinds of measures that the CSAC and the NQF Board is thrilled with.

So, they also suggest if you are thinking about teaching or counseling kinds of measures, they should be evaluated from the

patient perspective. So, not necessarily so much did you teach, but perhaps did the patient really understand what you taught might be another way to think about it. And I believe in Phase I you did consider an education measure.

Okay. Next slide.

Just a reminder. Again, scientific acceptability has two major subcriteria, reliability and validity. Both reliability and validity specifications are very important. The measure specifications are what you will be thinking about. For validity, you will really be thinking about how the specifications line up with the evidence, okay, and then all these other things.

Next slide.

Evaluation of testing, again, this only comes through with four of our measures, but this is just a slide reminding you that measures can be tested at data element level

or the measure score level. So, when you are thinking about testing and testing results, you have to think about: what were the results themselves? What level was testing done? When testing was done, was an appropriate method used, appropriate sample sizes, that sort of thing? That is the scope of the testing. And again, finally, the results of the testing. So, there are kind of a lot of moving parts when it comes to evaluating testing results.

Next slide.

And again, this is just to show you the scales that you will be using to rate validity and reliability for measures that have been tested. Again, just a reminder to get a rating of high in both cases for reliability and validity we would expect to see testing done at both the data-element and the measure-score level. So, if the testing result is a phenomenal result, I mean it is just really pristine, that is not enough to

give it a high. It needs to be tested at both levels. And then, if it is pristine, then it would get a high.

Next slide.

This is just the remainder of the scale.

And next slide.

We have a decision logic table that helps figure out if something passes reliability and validity and, therefore, the scientific acceptability. So, basically, again, the measures would need to have a high or a moderate on both reliability and validity to pass scientific acceptability.

Okay. Next slide.

The CSAC also offered some guidance around testing. One of the things that they suggested is you have to think about the impact of missing data. You shouldn't just make those exclusions when you are developing a measure.

Exclusions should be evidence-

based. Measures need to have the broadest applicability possible in terms of population settings as well as some analysis. And also, avoid measures where improvement decreases the denominator. Again, we won't focus too much on that guidance because I don't think it is really relevant for today's measures.

Next slide.

Usability, that is the extent to which intended audiences can understand the results of the measure and find them useful for decisionmaking. We ask you to think about public reporting as well as internal quality improvement efforts.

Okay. Next slide.

Feasibility is the extent to which data are readily available, retrievable, and easily implemented. So, it really gets a lot to data burden and that sort of thing.

Next slide.

And we can stop here. We don't have to talk about this right now.

Potentially, tomorrow we may need to talk a little bit about competing and related measures.

So, with that very brief introduction, let me see if there are any questions about what you will be doing today. Again, you have your four-pager in front of you. So, please refer to that if you forget what the scales are, and we will try to be putting scales and that sort of thing up on the screen as well when it comes time to vote.

DR. BURSTIN: Just one brief addition, and this is, I think, the first time you have seen untested measures. We have a strong preference for tested measures. But when there are clear programs that are going to be using these measures in the short-term, you want to have the chance to evaluate them, even if they are not tested. They have got to get their testing results done and in within 12 months and have a clear plan of how they are going to do that.

But, at the same time, since you can't really look at reliability and validity, it is very important to still look at the precision of the specifications. Do you believe the specifications are precise enough that they can logically be reliably collected, even if you don't have that testing data?

CO-CHAIR KNOWLTON: When we were talking as we were preparing for this session, I asked you whether we should talk about some of the criteria, some overarching criteria that seems to be at issue. Certainly, all of you who have participated in the Working Groups know that there have been issues over evidence. And I wondered if it would be helpful to have, which I would like you to guide, would it be helpful to have a discussion of how we all feel about that, because it was very much a repetitive issue n our discussions in the working group?

But we can apply these criteria on each individual one, but I wondered if it

would be helpful to talk about it first, the issue of clear and present evidence. Some of these are time-limited endorsements because the evidence isn't --

MS. JOHNSON: Let me make sure that everybody understands, when it comes to these untested measures, you will be considering is whether you will recommend them or not for endorsement. But, as Helen said, that would just be what we call time-limited endorsement. It would be for 12 months. During that 12-month time, the developer should be testing the measure, and then they would bring it back to us and we would evaluate the results.

But, basically, the non-tested measures just means, when you are doing the scientific acceptability criterion, you are not going to be looking at all those different things under reliability and validity. You will pretty much be focusing only on the measure specifications, again, the precision

of them and how they line up with the evidence.

Okay. Having no testing and being up for potential time-limited endorsement has nothing at all to do with evidence. So, they are not getting a pass, if you will, on having to show impact, high impact, having strong evidence base and having opportunity for improvement. I guess that is more to the first issues that we wanted to make sure that we clarified.

Does that make sense, Dave?

CO-CHAIR KNOWLTON: But even in a time-limited endorsement there has to be evidence, and it has to be specified. I think David actually made this point during one of our calls. There has to be very clear evidence; it has to be specified, and we have to be able to understand it --

MS. JOHNSON: Yes.

CO-CHAIR KNOWLTON: -- and how it is applied.

MS. JOHNSON: Yes.

CO-CHAIR KNOWLTON: Because that seems to have been a repetitive question.

DR. BURSTIN: To get to be a time-limited-endorsed measure, it has to meet every single criteria for any other measure with the exception of the fact that it has not been tested for reliability and validity. That is all. There is no separate bar. It is literally exactly the same with the exception of reliability and validity.

And in that case, what you are really looking at is the precision of the specifications and how comfortable you feel that, as this goes out into practice, in advance of having it tested, that that is going to be likely reliably in the 12-month period while we await testing results.

MEMBER COONEY: The common issue that came up during our discussion, you know, you have that slide that showed the connection between the process and the outcome and the

different levels of importance. That one.

Because a number of the measures are assessment measures. What I found lacking in many of what we reviewed was the connection to the outcome. Because it applies to so many measures, as David said, is it possible to talk a little bit about that, I mean the importance of that tie? Because the assessment seems useful, valuable. We should do it. But I didn't find the tie to the outcomes, and it seems like we are supposed to have a tie to the outcomes. So, could we discuss that aspect of it a bit?

DR. BURSTIN: Yes. So, in general, it is an excellent question, and this comes up a lot whenever we get a batch of assessment measures, which a lot of these are.

So, in general, the preference would be, if you are going to have process measures, then they have got to be as proximal to the outcome as possible. And certainly, assessment measures are usually pretty distal.

They are the first step in that pathway towards getting to an outcome.

So, I think the only time we have seen assessment measures come forward is when it is a relatively-new area, for example, where there hasn't been a lot of measurement done to date, where there is significant gaps in even doing the assessment.

And so, one question might be -- and this is where you guys may need to invoke that exception, and it is an exception; it is not a pathway; it is truly just an exception -- where you really look at that measure and you think, boy, I know enough about this topical area to know that 70 percent of the time clinicians aren't even doing that. It is so important to get this on the sort of measurement radar screen that we still think putting it forward with an exception, the benefits would so exceed the risks, that it is not so issue, you know, in your world, Gail.

For example, on our Palliative

Care Committee there was a measure that looked at whether somebody was offered spiritual services. Again, one of those things where it sort of seemed intuitive, like, of course, we want to get on this path, even though we were really pretty far away from a lot of the outcomes we were really interested in. But the Committee felt strongly that was one place where you would potentially invoke that exception.

MEMBER KAPLITT: So, all I would suggest is that the recommendation that was made at the beginning that we start each criteria with evidence I think really speaks to the point, because of the fact that that clearly is the major issue.

My concern about having an extended general discussion right now is that we are going to wind up back doing the exact same thing because each one has their own evidence issues. There may be overarching themes, right, which is evidence important or

not. And there may have been confusion, I think, before the calls by some people as to how evidence fit in versus reliability and validity testing. And if that is still an issue, then that may be worth a general discussion.

But my concern is, if we start talking about specifics of each thing, that is the reason why I think -- because, then, if we go through evidence as the first discussion point for each thing, if it falls down there, then we have saved a lot of time. And then, you know, that's that.

And then, I would personally -- I think it is likely that when we discuss the first measure for a given area, if the evidence falls down, the rest of them will probably go fairly quickly, because just based on the initial view, similar themes for different measures in the same area kept coming up.

MEMBER SULLIVAN: I have one other

general question that came up on our Work Group call, and it came up several times with several measures. That is the issue, it goes to specifications. But there are a number of measures that seem to lump a lot of things into the measure. You were assessing a number of things that were related but were different.

And I think that is likely to come up looking at a lot of these. I just don't know how to address those. I think it would be difficult in terms of usability of a measure if you are lumping six things that a clinician is supposed to be assessing. I wonder if there is some guidance in general about how to look at those measures that assess multiple related things.

DR. BURSTIN: Actually, that is something we see pretty commonly. Actually, many of them become composites or all-or-none composites, but you should always do all of these. So, that is not very atypical.

The key is going to be do the specifications have enough precision that you can, in fact, walk through it? Now there are also at times, when you are looking at multiple things, one could make the argument reliability may take a hit. And that is something, even though you have testing results in front of you, I think it is a fair question to invoke, you know, depending on the complexity of the data collection, is that something necessarily that you think can be reliably collected, even in advance of testing.

Did you have a question?

CO-CHAIR TIRSCHWELL: Yes, I just had one point that I thought was probably relevant to a lot of measures. It goes back to your slide about the two things that the CSAC doesn't like in the measures. One of them was a checkbox measure.

DR. BURSTIN: Yes.

CO-CHAIR TIRSCHWELL: And it seems

like a lot of these assessment things are literally checkboxes in your clinical evaluation. Once you check them off, you are good to know. Who knows that it leads to different treatment, let alone outcomes later that are changing things?

So, I mean, I have had doubts about some of the evidence for a number of the measures, as I think a lot of people did. But, then, seeing that I think is something we will probably have to keep in mind for a lot of measures.

And what was the second CSAC thing?

MS. JOHNSON: It had to do with --

CO-CHAIR KNOWLTON: Counseling.

CO-CHAIR TIRSCHWELL: Oh, right.

CO-CHAIR KNOWLTON: From a patient perspective.

CO-CHAIR TIRSCHWELL: All right. You gave an example of that. So, that is it. Thanks.

DR. BURSTIN: You know, in general, the idea of having somebody say, "I counseled the patient," and it is the measure, as opposed to understanding from the patient that that was in any way a meaningful event, is difficult.

CO-CHAIR KNOWLTON: Maybe it is my denseness in this, but does there need, without an exception, does there need to be a clear link between a measure and a desired outcome?

DR. BURSTIN: Yes, there should be a clear link, and if you think it is important enough to do anyway, then you would need to invoke the exception.

CO-CHAIR KNOWLTON: Got it.

Anybody have any other questions? I don't see anybody else.

Oh, quickly, go ahead, Michael.

MEMBER KAPLITT: A procedural question. Based on the agenda, we are also at the end tomorrow re-explore that Phase I

measure. Not to make too many assumptions, but if the agenda moves along quicker than anticipated because of this evidence issue, is the developer available to do this or are we going to just be hanging around until 1:30?

DR. BURSTIN: We will look into that. I mean, certainly, I think they would be available tomorrow. I don't know that they planned to do it today.

MEMBER KAPLITT: But I kind of have a suspicion maybe that the agenda may be moving quicker than this one is.

DR. BURSTIN: And you guys all heard, it is only one measure now, the readmission measure. CMS has withdrawn the mortality measure due to the concerns about risk adjustment.

CO-CHAIR KNOWLTON: And no matter how fast we go today -- thank you for raising that, Michael -- no matter how fast we go today, we still need to do tomorrow because the measure developers will be here for those

scores tomorrow. So, we still have to do tomorrow.

MEMBER KAPLITT: Right, but --

CO-CHAIR KNOWLTON: But that last measure, you are right.

MEMBER KAPLITT: -- first thing in the morning --

CO-CHAIR KNOWLTON: That's right. That is a good point.

The other thing, on the other side of that is it is our understanding that we only have one person who has to leave before our scheduled conclusion time. If anybody's change and that changes, you need to let one of the Co-Chairs know.

And just a housekeeping measure, when you put your thing up so that we call on you, make sure we can see it. Some people put it so I can't see the name.

So, we are going to start off now, and you know the process. We are going to begin with a lead discussant on the issues.

We are into the first measure.

MS. JOHNSON: Sorry, I think we put out a final agenda. So, we are missing one little thing that we wanted to do.

CO-CHAIR TIRSCHWELL: A "final" final agenda.

MS. JOHNSON: Yes, a "final" final agenda.

(Laughter.)

CO-CHAIR KNOWLTON: I am not on the "final" final.

MS. JOHNSON: You have the almost-final one.

CO-CHAIR KNOWLTON: Almost final?

MS. JOHNSON: Sorry, Dave.

What we are going to ask our developers for the first three measures to do is we are going to give you about five minutes to just give us a general overview of your measures, just so we can get to know you a little bit.

So, if the folks from PQA want to

start, that would be great.

MS. KUHLE: Good morning.

That was loud. Okay, is that better?

I am not sure it is best to go first in the morning, especially with a measure. You feel like you have got all the time to really scrutinize it.

Let me give you a little history of PQA. It is a consensus-based, multi-stakeholder alliance focused on initiatives to improve the quality of medication use. So, it is a little bit of a different measure that you are going to look at because it really does develop just using prescription claims data and then, of course, diagnosis data.

The members of PQA are diverse. We have pharmacist professional associations. We have federal agencies. We have health plans. We have academic institutions. We have pharmacists and chain pharmacies and independent pharmacists and consumer advocacy

organizations.

The measure development process occurs through Work Groups, and the Work Groups are comprised of representatives from all of our member organizations. So, again, they have diverse backgrounds and expertise.

This measure was initiated last year, 2011, in a Work Group called the Overuse Work Group. This Work Group wanted to look at this measure because of the growing evidence of poor outcomes for patients with dementia that were using antipsychotics and our understanding, as pharmacists, that antipsychotics are often overused.

This year, the Mental Health Work Group looked at this measure concept and further reviewed and revised it with their expertise.

And then, finally, the measure concept is reviewed by a quality metrics expert panel. That is a group of individuals with really specific expertise in the area of

prescription claims data, but also quality measurement and outcomes research.

And then, finally, this expert panel did review the testing. We have some limited testing of this measure. And then, the measure was brought forward to our full membership for endorsement. And that was last June.

So, that is my introduction. I hope that helps you understand where this measure came from.

Karen, I would also ask, do we have anyone on the phone? Because my colleague, Dr. David Nau, was going to call in.

MS. JOHNSON: I know I heard someone on the phone.

Dr. Nau, are you on the phone?

DR. NAU: Yes, I'm here.

MS. JOHNSON: Okay. Great.

DR. NAU: All I would add to what Julie mentioned is that we did have a

combination of physicians, pharmacists, nurses, and others that weighed-in on the development process and also made sure to test it.

But the institute that it was primary designed to evaluate was Medicare clients. So, we do have that testing evidence that is in the submission form that you have all evaluated.

We believe this is an important area that is very relevant for patients in the Medicare program and helps to give that population-level perspective on the use of these medications in patients with dementia.

MS. JOHNSON: Thank you very much.

And how about the folks from AMDA? Would you like to give us a brief overview of your measures?

MS. VANCE: Hi. I'm Jackie Vance. I am with the American Medical Directors Association. Our Association represents professionals who care for frail elders in the

long-term care continuum. So, our patient base is mostly in the nursing home setting, where the average age of the patient is 85 years old. Our measure is designed for the nursing home setting.

According 2012 Alzheimer's Association facts and figures data, the prevalence rate of Alzheimer's disease by the age of 85 is 47 percent. In 2011, more than 5.3 million American had Alzheimer's disease, while 2 million went undiagnosed. In 2009, 68 percent of nursing home patients had some form of cognitive impairment, 47 percent in the moderate-to-severe stage. Yet, in 2011, we looked at Medicare claims data from the MDS assessment, which I will explain what that is in a moment, and it showed that only 47 percent of those nursing home patients had an actual documentation in the medical record of having dementia.

According to a U.S. Preventive Service Task Force systematic evidence review,

it showed that 50 percent of patients with dementia have never been diagnosed by a physician at all.

Then, we noticed that in 1992 HCPR convened a panel of experts to develop a guideline on screening for Alzheimer's disease and related dementias, and I will quote them, "Failure to diagnose dementia can result in needless and possible harmful treatment and needless healthcare expenditures. We put out a lot more evidence, and we put that within our submission on that.

According to evidence such as an HCPR Guideline Overview No. 19, the correct diagnosis of dementia can prevent costly and inappropriate treatment. It all shows that awareness of dementia allows the clinicians to provide a prognosis and expectations, realistic expectations, and allows for things like improved pain detection, weight-loss intervention, elopement prevention, and other appropriate care.

This measure allows us to take advantage of what is unique to the long-term setting. That is several things that many of you might not be aware of. We have something called a minimum dataset assessment that is done for every person that is admitted to a nursing facility. It is licensed under Medicare or Medicaid. The MDS assessment was updated in 2010. It is a validated tool. I provided that evidence as well and all those validation studies in the submission.

One of the reasons why the MDS assessment was developed is nursing homes are mostly staffed by LPNs or LVNs who do not really understand or are not taught assessment techniques. They can evaluate; they don't assess.

And also, in the nursing home setting, unlike the hospital setting, it was a rude awakening for me. I started out in acute care and then moved to long-term care. You don't have physicians living in their

nursing home. They come every 30 days for the first 60 days of the person's life in the nursing home. It is a federally-regulated visit they must make. And then, they come every 60 days thereafter unless they need what is called a medically-necessary visit.

The MDS assessment triggers things -- it is actually called a care area assessment -- that will let the nursing staff know that something is going on enough with the resident that helps trigger that they need to let the practitioner know to come in and make that medically-necessary visit.

So, what we look for within this measure is looking for consistence, a brief interview, mental status assessment, that will give a certain score. That is certainly going to trigger that this person will most likely have severe dementia, but there hasn't been a diagnosis of dementia because the nurse can't make that diagnosis. That diagnosis is transcribed from the medical record onto the

MDS once that physician has made that diagnosis.

And then, an appropriate care plan or treatment plan can be put in place. We have evidence. We know that you have very negative outcomes, health outcomes, on healthcare cost outcomes when you don't have that appropriate diagnosis. You don't have advanced directives in place. You have things that happen, and I have seen it in my career, just these persons being sent back and forth to the hospital with futile healthcare treatments, expenditures, and you don't have these appropriate treatments in place that help maintain function, that help maintain whatever cognition that you can maintain.

So, what we are looking for is an assessment at least to a process and outcomes that you know will happen by what you will see happen within the future MDS documentation.

Thank you.

CO-CHAIR KNOWLTON: Thank you.

Salina, welcome. We want around and did a disclosure. Introduce yourself and please disclose any conflicts you have.

MEMBER WADDY: Sure. I am Salina Waddy from the National Institutes of Health, and I have no disclosures other than my job.

CO-CHAIR KNOWLTON: Okay. Now we are back to the agenda. Am I on the right one now?

So, we are going to begin with you, Gwen, on the Measure 2111, if you could take the lead of the discussion, please.

MEMBER BUHR: So, this measure is an psychotic use in persons with dementia. It is from the Pharmacy Quality Alliance. It is measuring the percentage of individuals 65 years of age and older with dementia who are receiving an antipsychotic medication without evidence of a psychotic disorder or related condition.

It is defining dementia as a diagnosis of dementia or being prescribed a

dementia medication. The people with psychoses are the ones with schizophrenia, bipolar, Huntington's, and Tourette's syndrome. Those are the ones that are having a psychotic-disorder-related condition where you could be on an antipsychotic.

And so, that is the introduction. On to the evidence, on our Work Group call we felt like there was quite a bit of evidence supporting the measure, that this was a process measure fairly proximal to the outcome, that it was prescribing a medication, and there are a lot of people with dementia. There is quite a lot of evidence that a lot of people with dementia are prescribed antipsychotic medications, and that antipsychotic medications in patients with dementia can result in negative outcomes, cardiovascular bad outcomes, and mortality. So, that was what I have to say about that.

CO-CHAIR TIRSCHWELL: So, I guess I had a question. I think somewhere in all of

this paperwork they mention that they don't expect the rate to be zero because there is some appropriate use of antipsychotics in patients with dementia.

But, then, my question comes down to, is dementia stage related to appropriate use? In other words, does it become more appropriate and more frequently used in patients with more severe dementia? And if that is the case, then does not this measure need to be risk-stratified, so that whatever prescription plan is being evaluated on this use is appropriately compared based on the types of dementia patients, more severe or less severe, that are in their particular plan?

CO-CHAIR KNOWLTON: Gail?

MEMBER COONEY: I deal pretty much only with patients with end-stage dementia, and I am unaware of any stratification of outcomes of the data that exists. It seems reasonable that, if you are increasing your

risk of stroke and death, then this risk of stroke and death should be less if you are near the end of life, but I don't believe that anyone has ever looked at that.

CO-CHAIR TIRSCHWELL: But that wasn't quite my question. I think, given any particular severity measure, it would be better not to be on it than on it. But if you are at a severe stage and it just becomes appropriate in 20 percent of the cases, and that is the minimum that you can get away with because it is really necessary versus mild dementia where you really can get away with 10 percent of people needing antipsychotics, you can't compare the health plan that has more severe to the health plan that has less severe and think that you are judging quality on that, when they are really both sort of at an appropriate level of treatment.

CO-CHAIR KNOWLTON: Jocelyn?

MEMBER J. BAUTISTA: Yes, I am not an expert in this field, but I did do a quick

literature search the other day, and I thought I did see a recent paper. So, the premise is that antipsychotic use increases mortality. The paper looked at in dementia patients that increased mortality was very much related to dementia severity, which I think is exactly what you are asking, right? So, those with the higher mortality were those with the more severe dementia. And this antipsychotic use may just be sort of --

CO-CHAIR TIRSCHWELL: But the antipsychotic use may be even an independent predictor of severity, and it is still problem for comparing the more severe group to a less severe group overall. You want to minimize it. I am not arguing with that at all. But what the appropriate rate is won't be the same for the healthcare plan that takes care of the more severe patients, is all I am saying.

And so, I just wonder whether certain healthcare plans -- and this is specified for a healthcare plan level, and it

is hard, knowing who is covered by different insurance, it is hard to believe there is not different populations of dementia patient being cared for by these different health plans.

CO-CHAIR KNOWLTON: Peter?

DR. NAU: This is David Nau from PQA.

CO-CHAIR KNOWLTON: Hold on. Hold on for a minute, David.

Peter?

MEMBER SCHMIDT: So, I am interested that Parkinson's disease is not included in here. Many movement disorder neurologists will tell you that, if the phone rings in the middle night and they pick it up, they say, "Clozapine." I don't want John Duda to be accused of poor-quality care for his patients.

So, is there a reason that Parkinson's disease is not included in the --

CO-CHAIR TIRSCHWELL: So, that is

a good point, but it is a specification issue. So, let's revisit that later, if we get to the specifications.

CO-CHAIR KNOWLTON: I have a concern here as well similar to David's. If we expect the number not to be zero in a measure like this, but we don't stratify how we know that it is zero, then one group of patients could be with one provider and all meeting that criteria. And so, there is no precision to the measure at all. So, it doesn't tell you anything.

I mean, I don't understand that clinically perhaps as well as you actual clinicians do, but I understand it logically, that it doesn't make sense, how we could differentiate the measures. That is my problem with this measure.

Ramon?

MEMBER R. BAUTISTA: We are required in the DSM-IV diagnosis of all these side conditions before we can even prescribe

an antipsychotic to demented patients then? Is that the idea here? Is there the DSM-IV criteria? I mean, what percent of general doctors out there know how to use the DSM-IV criteria for that matter?

CO-CHAIR KNOWLTON: Gwendolyn?

MEMBER BUHR: Well, I think in response to that, it is based on the diagnosis codes. So, it is going to be administrative data in that way.

And about the evidence and the different severities of dementia, I think the bulk of the evidence is just meta-analyses of the major trials of antipsychotics. You can't break it down by severity of dementia because the individual trials themselves are not showing mortality. But once you lump them all together, you have enough power to show mortality. So, you can't break them down by severity.

There is going to be cohort or other kinds of studies that might be able to

try to break it down by severity, but I think the evidence is more lumped together in meta-analysis. And if you just have the patient population, I mean, you are going to prescribe an antipsychotic for behavioral and psychological symptoms of dementia, which are happening at various stages in various ways.

And so, in the mild and moderate stages, you are going to have different things that you might prescribe an antipsychotic for than in the severe stages. And so, I just think that I don't know about the utility of breaking it down and risk-stratifying.

CO-CHAIR TIRSCHWELL: So, in direct opposition to my role as Chairman, I have confused the issue by raising the risk adjustment when that should be part of the specifications as well.

(Laughter.)

And so, we should probably table that and just talk about the evidence, as you, I think, described, Gwen, linking excessive

use of these antipsychotics to the outcomes of increased mortality.

Jocelyn?

MEMBER J. BAUTISTA: But I think it is still an evidence issue. If the premise is that use of antipsychotics is an independent predictor of mortality, increases risk of mortality, completely separate from severity of disease, has that been shown clearly in the evidence?

MEMBER BUHR: I think so.

MEMBER J. BAUTISTA: Didn't you just say it hasn't been adjusted for severity?

MEMBER BUHR: But if you have a randomized controlled trial of an antipsychotic with a person with dementia, and you have got various stages of dementia, then you can lump all those -- I mean, it seems like that it is, I don't know --

CO-CHAIR KNOWLTON: It appears, though, Gwen, it appears squishy when you say that the developer says we don't expect the

rate ever to be zero.

MEMBER BUHR: Well, the reason for that is because this is a difficult problem. Patients with dementia with behavioral and psychological symptoms, there are not good treatments. A patient may have some behavioral symptoms that are putting them at a danger to themselves or others, and you have to take the risk of higher mortality because the antipsychotic is the only drug that has much evidence behind it to improve their behavioral symptoms. So, you may be able to treat their behavioral symptoms and make them not be a danger to themselves or others, but they may have a negative cardiovascular or they may die sooner.

And families and patients -- well, patients really can't choose at that point -- but families are willing to take that risk because the patient is having such a poor quality of life, and they are a danger to themselves, a danger to others. They can't

live in their nursing home because they are going to hurt the staff or hurt themselves.

And so, you take the risk of higher mortality. And that is why it can never be zero.

CO-CHAIR KNOWLTON: But those, then, become extenuating circumstances and confounding variables that take a particular measure and say this measure isn't providing new, meaningful information to someone because the practitioner says, well, clinically, I had to make a tough call here.

MEMBER BUHR: Yes.

CO-CHAIR KNOWLTON: But the measure is trying to say what is the appropriate call, and we don't have evidence of the appropriate call here. I mean, you guys are the clinicians, but you are going to paint --

MEMBER BUHR: Isn't that really validity, though?

CO-CHAIR KNOWLTON: You are going

to paint the clinician into a box, aren't you?

DR. BURSTIN: Again, keep in mind it is a health-plan-level measure. So, you are not really at the clinician level as a starting point.

I think one of the questions that I would be curious to have David or somebody respond to this is the point you raised earlier about are there differences by types of health plans. And for example, many of the NCQA measures on the health plan level are stratified by type of health plan, Medicare, Medicaid, commercial. I mean, maybe that is one approach, but I agree that is more risk adjustment.

But I think, David, it is actually not that dissimilar to other measures we have had where some things just can't be zero, but we would like them to be low. One example is episiotomy after birth. I mean, it is not something that is ever going to be zero. You would like it to be low.

And again, by having the comparisons across providers, it is very useful for people to understand and drive improvement. But, again, getting to zero doesn't make sense.

DR. NAU: This is David.

I think you said a key point there. But, also, with regard to risk stratification, certainly with this measure there can be the same sort of very simple stratification as used with many of the claims-based measures, which are just to segment the type of plan generally by Medicare, Medicaid, et cetera.

But, additionally, in trying to do risk stratification based on severity of dementia, that is just not possible to do with claims because the ICD-9 codes don't allow for that sort of detailed nuance to be put into the claims. So, it is only capable of identifying those who have been diagnosed with dementia.

But when you look across the large or even moderate-sized health plans, you are talking about thousands or hundreds of thousands of patients. In general, the distribution across different risk strata are similar across the plans. And so, it doesn't appear as though it would greatly bias one plan over another from that regard. In fact, we found fairly consistent results across a few of the plans we did evaluate this in.

However, the underlying clinical evidence, the studies, the randomized controlled trials that were conducted did account in many cases for the severity of the patient and still found a significant elevated risk of antipsycholotic use in the patients who had dementia.

CO-CHAIR KNOWLTON: Very good.

A.M.?

MEMBER BARRETT: So, I want to echo the concern about the evidence linking this process measure to outcomes, not just

because of my being unclear about the proper percentage of people who would appropriately receive antipsychotics for behavioral and other symptoms of dementia, chronic symptoms, I assume, but also please educate me, Committee, if I am correct, but agitated delirium also can be a risk factor both for mortality and an indication for antipsychotic drug use in dementia. I am not sure that the meta-analyses that were done took into account that influence.

CO-CHAIR KNOWLTON: Jordan?

MEMBER EISENSTOCK: I am not sure if I am out of place now. That is why my card has been up and down after what you said, David.

(Laughter.)

But I think that your point is incredibly important here. That is that, as time goes on, the big global risk versus benefit factors change for these patients. This is sort of dovetailing what A.M. just

said as well.

I am just trying to think of some way to reconcile it because I think everybody's point is really right on, but it is how to bring it all together and make it fit.

As these patients age and they go into more moderate or severe levels of dementia, we find that the overall global risk factors sometimes predispose to providing or prescribing the antipsychotics, and that is because they are truly psychotic and there are no other treatments.

I think the intent of this measure was to avoid using antipsychotics in patients that could otherwise be treated for aggression or agitation in other less risky ways. So, in playing upon that intent, one way perhaps to reconcile this would be to just another diagnosis to the numerator statement. I know we don't like "not otherwise specified" very often, but if it was psychosis or psychotic

disorder not otherwise specified, it might include more patients who are being treated properly with antipsychotics and help to reconcile the issues that we are discussing right now.

CO-CHAIR KNOWLTON: Risha?

MEMBER GIDWANI: I think the last speaker did a good job of summarizing that. I sort of will just build on that.

I think Gwen makes a good point about the patients being a risk to themselves or others, and therefore, use of antipsychotics is appropriate. In fact, on page 9 of our documentation, the guideline for the American Geriatric Society actually says the recommendation is to avoid use for behavioral problems of dementia "unless non-pharmacological options have failed and patient is a threat to self or others."

So, I am wondering if the developer can address why their numerator statement departs from this guideline.

CO-CHAIR KNOWLTON: Does the developer want to answer Risha's question?

MS. KUHLE: I am not sure if Dave is on the line.

Two things. If the idea is to make sure that patients that are a danger to others can receive this medication and not be counted in the numerator, and we can try to do that with an ICD-9 code, absolutely.

But it is my understanding that patients who have agitation, have behavioral symptoms that can be otherwise managed with non-pharmacological treatment, shouldn't receive antipsychotics. And that is really what this measure is trying to get at.

And the idea that, when there is harmful behavior, absolutely, that is when we want them to be treated with an antipsychotic if that is the last choice. But it is also my understanding that patients have acute need, and then it is not as if, once they become aggressive, they always will remain

aggressive. They might have outbursts where they need acute treatment and then they don't. And what we don't want to see is that patient stay on an antipsychotic forever.

One of the criteria for this is longer than 30 days' supply of the antipsychotic. So, we are looking for more than just an acute use.

I hope that helps answer your question a little bit.

CO-CHAIR KNOWLTON: You can follow up, Risha. Go ahead.

MEMBER GIDWANI: Well, I am not a clinician. So, I can't really speak to what the appropriate use of the antipsychotics is. But if the clinicians in the room are able to talk about whether a lower-than-30-day supply would be considered appropriate for treating this threat to one's self or others, and then, beyond that, we would say that, yes, this is definitively an inappropriate use of antipsychotics, that would help me in deciding

whether this is a valid measure.

CO-CHAIR KNOWLTON: I will let a clinician answer Risha's question. And then, I am going to ask if we focus on evidence and ask if you have your card up for something on other issues -- let's get to a vote on the evidence. I don't know whether, Gwen, since you did this discussion, would you want to answer?

MEMBER BUHR: So, I think that with respect to the evidence and the measure, and whether it is appropriate to use an antipsychotic, it is that even if the person has -- the evidence says that a person who has dementia, regardless of anything else, is at risk for mortality and cardiovascular outcomes. And so, you don't want to use an antipsychotic for anything, but sometimes you have to because non-pharmacologic measures aren't working and other safer medications aren't working. And so, then, you take the risk.

And so, does that help at all with your question?

MEMBER GIDWANI: I think it is just more the 30-day issue.

MEMBER BUHR: Okay, the 30-day issue, yes. So, I guess that you would simultaneously, as you are using an antipsychotic, be using non-pharmacologic measures. And those may work over time or you may figure out one that does work. And then, you may simultaneously use other medications that take longer to work, and so, then, after 30 days, be able to get rid of the antipsychotic.

And people with dementia, you know, they wax and wane in their symptoms. And so, you prescribe one, and then the rules say that in a nursing home at least you have to reduce the dose at certain intervals anyway.

CO-CHAIR KNOWLTON: John, on this issue?

MEMBER DUDA: So, one disclosure I forgot to mention is I wasn't here for part one and I was on a plane for my small group's conference call. So, I really have no idea what I am doing here.

(Laughter.)

CO-CHAIR KNOWLTON: That is a common feeling.

(Laughter.)

MEMBER DUDA: Like Jordan, I keep putting it up and down because I don't know when I am supposed to talk and when I am not.

But, as a clinician who takes care of patients with Parkinson's disease who get psychosis all the time, it is not at all uncommon -- so, one question, do they get a separate diagnosis of psychosis coded? Not necessarily, right? If you have PD and dementia and you put them on Seroquel for their psychosis -- I work in the VA, so coding isn't as much of a problem -- so, I don't know that we would necessarily be missing some that

way as well.

But, then, the 30-day issue, I have a patient with Parkinson's disease and dementia who is not going to get better, and he has a psychotic episode at one point. I am not likely to take him off, even if he is not psychotic because there have been good studies that show that this progresses. We will even treat people with kind of insight-retained hallucinations, because we know that it progresses to insight-unretained psychosis in the future, to try to prevent that.

So, my thoughts.

CO-CHAIR KNOWLTON: Thank you.

Ramon, on evidence?

MEMBER R. BAUTISTA: Yes. If I recall my medical school and residency training in psychiatry for a patient I had anyway, the sedative symptoms actually from Haldol, for example, work right away. But the psychosis symptoms, it takes weeks before they manifest.

So, to answer the question, I am not sure if the antipsychotic is going to cure your psychosis. It might help sedate you, but not take care of your psychosis. That takes a longer period of time. So, as far as I can remember, taking a one-month prescription of an antipsychotic doesn't help your psychosis.

CO-CHAIR KNOWLTON: Gail.

MEMBER COONEY: Real quick, one of the things I really like about this measure is that it doesn't have all those clinical exceptions, because I think there is very strong evidence that in this population these drugs should be avoided. And that is all this measure is really seeking to say.

CO-CHAIR KNOWLTON: Anything else on evidence?

(No response.)

Can we vote on evidence?

Okay. Can you open it up for us?

MS. THEBERGE: Okay. Before you vote, we have made a small change to how we

would ask you to think about the evidence. Basically, if you look at that slide, the two slides on the side, if you feel like you need to choose no for evidence, we would like you to try to distinguish for us whether it is "no" because the evidence is there but it doesn't meet the criteria or is it "no" because the evidence just didn't make it to the submission form. Does that make sense?

So, if it is yes, you do nothing different. You just vote yes. Okay? But if it is no, tell us again -- it is on this slide here -- evidence does not meet the guidelines or there is insufficient evidence presented for you to make that determination.

So, again, we are just trying to be -- and this will really come into play, I think, later on in the day -- we want to be very transparent about, if you vote things down on evidence, why exactly did that happen? Okay?

Any questions before we go to this

vote?

(No response.)

Okay.

CO-CHAIR KNOWLTON: Okay. Can we open it up now? We are ready to vote.

You should be pointing toward Suzanne. You can do it as many times as you want; it will only record once.

(Vote taken.)

MS. THEBERGE: Sixteen yes; 2, no, evidence does not meet guidance, and 5, no, insufficient information submitted on evidence.

CO-CHAIR KNOWLTON: Okay. It passes on evidence.

Back to you, Gwen, impact.

MEMBER BUHR: Okay. So, impact, our Work Group felt like it had a high impact because there is a lot of people with dementia, and a lot of people with dementia are being prescribed antipsychotics. So, high impact.

CO-CHAIR KNOWLTON: Questions?

(No response.)

We can vote.

(Vote taken.)

MS. THEBERGE: We need two more responses.

Twenty high, 2 moderate, 1 low.

CO-CHAIR KNOWLTON: Okay. Opportunity for improvement is next.

MEMBER BUHR: Okay. So, from the information they submitted, between 14 and 16 percent of the Medicare Advantage patients with dementia are receiving antipsychotics. And even if we don't want the number to be zero, we felt like there was a lot of opportunity for improvement with that.

CO-CHAIR KNOWLTON: Okay. Hold on for a minute.

MS. JOHNSON: Again, this is a little bit new, but if you have any discussion at all about performance gap, I would also ask, is there any flavor from the Committee

that this may reflect a disparity-sensitive issue? And it is okay to say no. But if you know of any disparities that might be around this measure, we would like to understand that.

CO-CHAIR KNOWLTON: Gail, are you waiting to speak?

MEMBER COONEY: No.

CO-CHAIR KNOWLTON: But A.M. is.

MEMBER BARRETT: I think that there are a number of studies showing that people from minority and racial backgrounds, cultural and racial minority backgrounds, are managed differently with dementia and perhaps with less quality.

MS. JOHNSON: Okay. And do we have any idea at all about in terms of antipsychotic use? It is okay to say no.

(No response.)

Okay. Thank you.

CO-CHAIR KNOWLTON: Okay.

MEMBER BUHR: I think in the stuff

they submitted there was something about some nursing homes have a much higher rate of antipsychotic use than others, and that suggests there is some kind of a disparity, but it is not understood as to why, or whatever.

MS. JOHNSON: And again, just a reminder, disparities, even if overall performance was great or in this case really, really low, if it turned out that there were disparities kinds of things, like you were saying that some nursing homes maybe are not performing so well, that would be another indication to you that there is room for improvement.

CO-CHAIR KNOWLTON: Peter?

MEMBER SCHMIDT: So, in performance gap, we are talking performance gap versus the measure or performance gap versus the evidence? So, for example, does John have a performance gap versus the measure or a performance gap versus the evidence?

MS. JOHNSON: We would be looking at for the measure. So, in this case, they have told us that the rate of antipsychotic use is between 14 and 16 percent.

What would be even more interesting would be knowing what the distribution of that would be to see, is it kind of fairly low, fairly uniform or really different? So, what we really want to see here is statistics about the measure itself.

CO-CHAIR KNOWLTON: Anybody else?

(No response.)

Okay, Suzanne.

(Vote taken.)

MS. THEBERGE: We still need two more.

All right. Eleven high, 11 moderate, 1 insufficient.

CO-CHAIR KNOWLTON: Okay. We are on to acceptability, scientific acceptability.

Gwen?

MEMBER BUHR: So, is this where

reliability -- okay, all right. I got it. Shall I talk about reliability separately?

MS. JOHNSON: Yes.

MEMBER BUHR: Okay. So, reliability, we thought that the measure was specified in a way that you could reliably measure the same people every time. That would be reliability. So, we didn't have concerns about reliability.

We did have some concerns about validity, but I shouldn't talk about that right now, right?

MS. JOHNSON: We will vote separately on reliability and validity. But this is where we would talk about precise specifications.

So, Peter's question about Parkinson's patients being included in the specifications would probably come up right about now. So, maybe we might want Peter to go ahead and just ask that again.

CO-CHAIR KNOWLTON: Go ahead,

Peter.

MEMBER SCHMIDT: I have a question about the reliability of the specification. And there are RCTs on antipsychotic in Parkinson's disease. You know, there is plenty of evidence.

I was interested to note that there is one study cited in the evidence section about Parkinson's disease, and it is about the correlation between antipsychotic use and hip fracture, which is common in Parkinson's disease anyway. But I am sure that there is a higher prevalence of hip fracture in people taking antipsychotics. I don't doubt that.

But that should not be imputed to indicate that antipsychotics are contraindicated in Parkinson's disease, which I think is having that be the only evidence around Parkinson's disease included in the study, is the implication, and I don't agree with that.

MEMBER DUDA: Yes, I mean, I think that that should be a consensus statement that is not difficult to reach consensus on. We obviously use antipsychotics to a great deal in patients with Parkinson's disease.

But I think it gets back to what Jordan was saying. If there were some way to identify the patients who had psychosis, instead of just agitation or whatever, you know, something else that is less indicated, then we could get around this. But, like I said, I am not sure that it is acceptable to expect every patient with Parkinson's disease who is put on Seroquel to have a separate code for psychosis, NOS, or related to a separate condition, you know.

MEMBER KAPLITT: In Huntington's we do, though.

MEMBER DUDA: Yes, but that, I think, is because we recognize that we don't use it for psychosis in that case, right? We use it for chorea. So, that is a different

reason.

MEMBER BUHR: So, this is a question -- maybe you guys who do Parkinson's know the answer -- but when patients have dementia from Parkinson's or Lewy body dementia, were they excluded from the trials of people with dementia and behaviors that are -- and so, there is not evidence that patients with Parkinson's disease who have dementia and take antipsychotics are at increased risk of mortality? Is there not evidence there? And so, we would want to limit or try not to use antipsychotics in patients with dementia in Parkinson's or not?

MEMBER DUDA: Yes. So, with the potential conflict that I have worked with the Lewy Body Dementia Association, who has advocated strongly in this matter for patients with Lewy body dementia, when these results came out, there was a big backlash thing. Yes, maybe there is increased mortality, but, obviously, the benefits outweigh the risks in

most of these patients.

As far as whether or not they were included in these studies, obviously, there were some patients with Parkinson's disease dementia because you can't clinically separate out Alzheimer's disease and Parkinson's disease with 100-percent certainty and specificity.

I don't know, I am not aware of any studies that specifically looked at patients with Parkinson's disease with neuroleptics and looked at mortality. I think there were some small studies in dementia with Lewy bodies, but they weren't as big as the ones in Alzheimer's, obviously.

CO-CHAIR KNOWLTON: Risha?

MEMBER GIDWANI: I had a question about the measure specifications. So, the patients, they looked at folks that had diagnosis codes for dementia and medication markers for dementia, and they saw that there were more patients identified when you have

both the diagnosis code for dementia and a drug marker for dementia. Therefore, one should be using both of these things to identify dementia patients. And that just doesn't sit properly with me.

If we are looking just at higher numbers of patients that we get from using both of these different ways of capturing them, we also have to look against their charts or some other gold standard to say, yes, these are the appropriate patients. Just we got more patients when we looked at a drug marker for dementia doesn't mean that that is the right way to capture those patients. That rests on the assumption that all of the medication use that is prescribed for dementia is appropriate, and I am wondering if there is any overprescribing of the dementia-related medications. And if that is the case, then do we need to be capturing more patients that we would be erroneously considering appropriate for inclusion in this measure?

CO-CHAIR KNOWLTON: Gail?

MEMBER COONEY: I think they addressed that somewhere in the data, that they actually looked at the underdiagnosis of dementia and using the prescribing of the cholinergics for that, cholinesterase inhibitors, whatever.

I also had a question because the ICD-9 codes that they include for dementia is much shorter than the ICD-9 codes used for the other dementia measures we are looking at. I don't know all my ICD-9 codes well enough to know why some are in and some aren't, but I had a question about that, too.

CO-CHAIR KNOWLTON: Michael?

MEMBER KAPLITT: So, I just think, to this point of Parkinson's disease or, frankly, anything else that could be excluded, you know, John's point is correct, which is that the reason that those specific things are being excluded is not because of the fact that psychosis doesn't matter in those diseases,

because the drugs are being used for a different reason.

I would argue with Parkinson's disease and other things -- and the other thing is that these measures are not just meant for experts, right? They are meant for the vast majority of general neurologists and internists and others who are treating the majority of these patients and probably using the majority of these medicines in these patients.

And so, I think the point here is that, if part of the consequence of this is it forces people to have to put in a psychosis code, if that is the rationale for giving that patient the antipsychotic, right -- if someone has Parkinson's and they believe that they have psychosis, I don't see the problem with forcing them to sort of put in a psychosis code and justify it, and maybe have to think about it for a second.

I mean, I understand all the

nuances of the issue, but, really, from the general standpoint I think, unless we can make an argument that there is something different about those patients that makes them a separate breed -- I see a lot of patients with Parkinson's disease who come from the general community, and a lot of them are on medicines that have not really been well-thought-out. So, I don't think it would be such a bad thing to require, and then that would meet the criteria here.

MEMBER SCHMIDT: So, I am actually going to kind of reverse myself. I agree with the last comment. There only are two antipsychotics that are generally considered safe for people with Parkinson's disease, and those aren't the ones that people with Parkinson's disease mostly get in a community setting.

Often -- I mean, John can tell me whether I am right about this -- but, often, psychosis can be managed by optimizing

existing medications and not adding on an additional one. And so, it probably is a good idea to have a general rule that antipsychotics should not be a "go-to" drug. You know, Haldol is terrible for people with Parkinson's disease and it is quite commonly prescribed by non-experts.

MEMBER DUDA: So, I guess I have a question to start out with. I mean, I was kind of thinking the same things that Mike was thinking. But is the purpose -- again, I am not quite sure what we are doing here -- is the purpose of this measure to guide future behavior of clinicians or to evaluate prior -- so, if we are using this, if we are saying, "Okay, Insurance Company, go out and use this to evaluate prior behavior," well, then it is not fair to expect those people to do things that we only are saying that they should be doing from this point forward.

CO-CHAIR KNOWLTON: Remember this is a health plan measure.

MEMBER DUDA: Right, that is what I am thinking. This is not a kind of individual practitioner measure. So, I think it is tricky to do that.

Not to muddy the waters any more, but he is right, there are problems with the prescribing of antipsychotics in Parkinson's disease that are systemic, and there are efforts to try to improve awareness than some are better than others. But there are also complications in Parkinson's disease. Some people use Clozapine for dyskinesia. It is a completely separate indication, and it is not an inappropriate indication. It actually has been approved by some consensus panel.

So, Parkinson's disease I think probably should be on that list for that reason, because you don't know if you are treating psychosis or dyskinesia, but, then, there should be some other way to pick up these other people with dementia who have psychosis.

CO-CHAIR KNOWLTON: I want to go back to Gail's question real quickly. Gail asked why was the list of dementia ICD-9 codes quite a bit shorter than some of our other measures. And we are not asking you to have necessarily the same list, but it does beg the question of why did you use these particular codes and not others.

So, maybe one of the developers, Dr. Nau, or the folks in the room, would care to answer that.

DR. NAU: Well, I have not looked at the other measures' list of ICD-9 codes. So, I can't speak specifically as to why they included certain diagnosis codes.

We did work with quite a few different experts, and we also looked at the studies, the epidemiological studies that have studied this issue of antipsychotic use in patients with dementia and tried to be judicious in what ICD-9s we included. And so, there is a lot of work back and forth and

refinement of our list, but I guess we would have to talk to the other measure developers about why they chose their lists the way they did.

MS. JOHNSON: Okay. So, just to rephrase, you had some experts weigh-in on the ones that you thought should be used, and you are pretty comfortable with that, at least for now? Talking about comparisons of lists would be something we could potentially do later, but I think it was just a question that came up.

CO-CHAIR KNOWLTON: Daniel?

MEMBER LABOVITZ: Yes, I wanted to just go back to your original point that this is, ultimately, an incredibly squishy measure. There is no way to know, even in a population, whether the use of the drug is appropriate or not. So, one healthcare plan might have a very high rate and it would all be perfect, and another healthcare plan might have a very low rate and have it be inappropriate.

That said, I can say from my own observation I think the general perception is that these medications are way overused. They are used for sleep. They are used for being mean. They are used for just keeping people quiet.

And I think, in the end, this is not dinging an individual provider. This is not Medicare coming after you and saying, "We're taking away 2 percent." This is really a chance to look at healthcare organizations in a broad swathe and say, "How are you doing with these drugs?"

That makes me much less inclined to turn the thumbscrews on issues that aren't -- for lack of precision. I think, in fact, the developers were thoughtful in making this imprecise.

MEMBER R. BAUTISTA: My concern why I think this might not be a reliable measure, though, is because, although it is not that hard to do the math and look at all

your CPT codes and ICD-9 codes and do the math and look, I am concerned about the fact that we may not even be correctly coding these things. In other words, we are not just measuring psychosis; we are measuring schizophrenia. We are not just measuring a sad person; we are measuring bipolar disease. These are four criteria, and I even doubt that the actual data that you actually put in, if you are not a psychiatrist, would be correct.

CO-CHAIR KNOWLTON: Gail?

MEMBER COONEY: I agree with Daniel that I think this measure is important, outweighs the concerns that are being raised. And also, I like its lack of specificity. I like the fact that it doesn't allow a lot of exclusions.

Even the narrowness of the diagnoses for bipolar and schizophrenia is useful because those things get tossed around without ever being correctly analyzed. So, I think it is important that we use those

diagnosis codes to exclude them, to make sure that we are not including people whose diagnosis was made randomly.

MEMBER R. BAUTISTA: The current way it is coded, though, here, if you were actually a non-psychiatrist coding the diagnosis, I would probably doubt that your diagnosis was correct in the first place. So, if you are using that as your current evidence for showing how great this measure is, how reliable might be, that in itself is a very invalid conclusion.

MEMBER SUDO: So, I think a lot of what we are talking about is validity, whether we are choosing the right population. That is validity.

So, the reliability is, can we every time get the same set of people? So, that is one point I wanted to make.

And the next thing is I am not sure we want to exclude psychoses because, I mean, it may be that Parkinson's disease is a

special thing, but psychosis, where you may prescribe an antipsychotic and in that sense it is appropriate, but the people who have psychoses and dementia and get an antipsychotic have higher mortality. So, we are trying, even though they have psychoses, to use other measures before the antipsychotic. And so, I don't think it would be right to exclude psychoses.

MEMBER VAN DE KAMP: I think Michael made a really good point earlier. I think it is sometimes easy to lose it in the group of such experts. There is a lot of, living in the nursing centers as much as I do and seeing the high usage, it is really a measure that just makes an awareness and I think has an opportunity to improve the quality of care for the general practitioner.

And that is, I would think, one of the goals of this. It isn't for the experts who really have the subtlety and the assessment skills. It is really for that very

often primary care physician who is just requested by the nursing center for behavior issues without really looking at all the pieces. So, I think it has a significant value and practical application within the healthcare environment.

MEMBER SCHMIDT: So, my question, I have a philosophical question, and that is, are we comfortable with a measure that would mark down the experts, where the true experts would get a lower score on this or a worse score?

Harvard Pilgrim has got a lot of experts in it. There are a lot of these centers -- there are a lot of health plans that have systematic referrals to expert care. We have seen this in -- you see this in health plans. There are some health plans that are based around academic medical centers or conglomerations of academic medical centers.

MS. JOHNSON: And just a reminder, this is not specified for clinicians. So,

this is specified for the pharmacy benefit plans. I think that would kind of get to your question, Peter.

MEMBER BUHR: I don't know that we can know whether the expert, like the Harvard plan would be worse or not, because maybe they have more resources. Maybe they will refer people for the different non-pharmacologic things. Maybe they have support groups. Maybe they have lots of educators to educate the families and the caregivers, because that is really where the evidence lies, in that the best treatments are educating caregivers about how to deal with these people. So, maybe Harvard has more resources in that way than another plan. I don't know that we would know that.

MS. KUHLE: I don't know if I can jump in real quick, but there is an old adage -- I'm sure you have all heard it -- that what isn't measured doesn't improve. And that was really the point of this measure, was to draw

attention to it.

Remember, these are all ambulatory patients, not necessarily just in nursing homes. They are a lot of dementia patients that are living at home, treated by family practice physicians.

Hoping that we can have an impact to improve performance is the whole goal of this measure.

MEMBER KAPLITT: I just want to clarify a point you made earlier when you were saying about maybe we shouldn't be excluding psychosis. So, then, I just want to make sure I understand. The suggestion would then be that the measure just be all patients who are on antipsychotics with dementia, period?

MEMBER BUHR: Yes.

MEMBER KAPLITT: I just want to make sure I understood what you were saying.

MEMBER BUHR: I mean, I didn't mean to not exclude the things that are already excluded, but those are very specific

diagnoses, Tourette's and Huntington's, and whatever. And they are not excluding people with dementia who have psychoses. They are not excluding that currently.

I was saying we were having some discussion about whether it should be, and I don't think that it should be because people with psychosis and dementia, while you may have to prescribe an antipsychotic to treat it, you try to treat it in lots of other ways. Those are the people with increased mortality.

MEMBER KAPLITT: But, I mean, I just think the concern there is that that is an indication for antipsychotic drugs, if they have a defined diagnosis of psychosis, and I think that is kind of a slippery slope at that point, because then you are basically saying to people -- I mean, I agree with you that you should try not to use them under certain circumstances, but now we are getting into a level of micromanagement that would concern me, because now you are saying plans are going

to have a problem if they prescribe an antipsychotic for a patient with psychosis because it is possible they didn't try enough other things first. I mean, that would concern me because that is an approved, appropriate indication for those drugs.

CO-CHAIR KNOWLTON: Let me ask if we can vote on reliability, get that out of the way, so that we are not here on Sunday.

So, let's put up the reliability vote, and please vote.

(Vote taken.)

MS. THEBERGE: Seven high, 12 moderate, 2 low, 2 insufficient.

CO-CHAIR KNOWLTON: So, this passes on reliability.

Validity, we have had that discussion.

John, I didn't mean to cut you off if you had another point.

Have we had enough discussion on validity? Okay, let's go to a vote.

MS. JOHNSON: Just real quick, is there any more discussion about risk adjustment or stratification?

DR. BURSTIN: And particularly of David Nau wants to respond to that because we never let him respond.

MS. JOHNSON: Oh, that's right.

Dr. Nau, this kind of goes back to one of the first things that we talked about with your measure. Did you have anything you wanted to add in terms of staging of dementia and stratification of the health plans? Stratification of health plans, yes.

(No response.)

MEMBER BUHR: So, could I make a comment about validity, because I did not make my validity comments earlier?

CO-CHAIR KNOWLTON: Sure.

MEMBER BUHR: Okay. So, in response to whatever Risha was saying about -- we did have a lot of discussion in our Work Group about the validity, because they are

measuring it by a diagnosis of dementia, and we know that dementia is way underdiagnosed. And they are measuring it by use of these medications, which we know that they are used sometimes inappropriately.

So, I have seen patients on these medications who don't have any signs or symptoms of dementia. In the stuff the developer presented to us, they say that they are used for traumatic brain injury, and Memantine is used for another indication. But they say that it is used rarely for those reasons. I don't know that we really know how rare it is.

By the way that they have told us that they have gathered their patients, they got prevalences of 5.3 and 7.2 percent. So, that is a much lower prevalence of dementia than is really thought to be the prevalence of dementia.

So, I think that, from our Work Group calls, our main concern of this measure

is, are we really gathering the patients with dementia with the way that they have specified it, knowing that that is the only way that they can specify it because it is claims data and pharmacy data, and whatever? We are not going to go and interview the patients and find the undiagnosed people, but that was our main concern.

CO-CHAIR KNOWLTON: Anybody else? Yes, go ahead, Jordan.

MEMBER EISENSTOCK: Just to sort of follow through with that, because that was one of my concerns in the Work Group call also. And I am going to try to be very diplomatic here because I like the intent of the measure, but I have big problems with both the numerator and the denominator in this measure.

And I just wanted to sort of put that out there because I do agree with what Gwen said. I think that it depends on how comfortable we are with the error that we know

is built in on both sides. Both the numerator and the denominator we know are not very perfect, and if we are okay with that is really what it comes down to with validity.

CO-CHAIR KNOWLTON: Any other comments?

(No response.)

I have one. I am concerned, as you are -- I am not the expert here; I am not the clinician here -- but, as was said earlier, if you don't measure it, the way I always said it was the only way to change something is to keep score. But you have to keep score in a way that people understand and is fair or people don't pay attention to the score. They say it doesn't matter; I can define it any way I want.

And that was my problem with this measure. I think that it can be really defined. It is squishy.

And I think I agree that we should have aspirational goals, but when they become

measures, we are asking people to be measured according to it, and I think that makes it difficult, from where I sit, more on the outside looking into this. But that is just my opinion.

Do we have any other stuff on validity?

Risha, I'm sorry, I didn't see it. Go ahead.

MEMBER GIDWANI: Just a brief comment. Yes, I have the same concerns about the sensitivity and the specificity of these codes and prescriptions to be able to capture this population.

Just from a measurement standpoint, this could be actually addressed by doing a chart review and getting a few hundred charts of patients that have a diagnosis of dementia and seeing what their codes were and their prescription claims, and then 200 patients that just have no diagnosis of dementia and seeing how many of them

actually really do have dementia that wasn't coded as such. So, there is actually a way to test the sensitivity and the specificity, but that wasn't done here.

CO-CHAIR KNOWLTON: Gwen?

MEMBER BUHR: I mean, one problem with looking for people in a chart review that don't have a diagnosis of dementia is, unless you specifically test them with different tests of dementia, you are not going to find the dementia, because there are lots of people who have dementia, but their doctor isn't really looking for it and he is just treating their hypertension, and whatever, and not asking them any memory questions, not asking them to draw a clock, or any test of dementia, and not uncovering the dementia. So, I don't know that you are going to fund the undiagnosed dementia population with a chart review.

MEMBER COONEY: The question about the ICD-9 codes are part of the denominator.

How does that enter into what I think about this measure? If I think they need to go back and standardize the ICD-9 codes, how does that affect my vote here?

MS. KUHLE: Can we say that we will do that? As a measure developer, we will work with the other measure developers to make sure that our codes are standard?

DR. BURSTIN: To me, it is a harmonization issue that we would address, depending on the dementia measures left standing, yes.

CO-CHAIR KNOWLTON: Risha?

MEMBER GIDWANI: Gwen, you make a good point. I think it would be hard to look at the correlation between patients that actually have dementia and it being documented in their chart.

What I am talking about is just the documentation of dementia in the chart versus the correlation with the ICD-9 billing codes. I have done some work in actually

looking at the correlation between what is written in clinical documentation and what the ICD-9 codes can capture, because the folks who are doing the actual billing operate within a very narrow purview and they can't interpret the clinical documentation. That is the component that I was really talking about.

CO-CHAIR KNOWLTON: Anybody else?

(No response.)

Okay, let's vote.

This is no validity.

(Vote taken.)

MS. THEBERGE: One high, 9 moderate, 12 low, 1 insufficient.

CO-CHAIR KNOWLTON: It did not pass on validity. Does that mean we stop here? We stop here. Okay. It has to pass on validity.

(Whereupon, the above-entitled matter went off the record at 10:55 a.m. and resumed at 11:16 a.m.)

CO-CHAIR KNOWLTON: Are we back?

Okay, we are moving on to the next measure, which is 2091, persistent indicators of dementia with other diagnoses, and it is Jocelyn, right? No.

MEMBER SUKO: No, Jolynn.

CO-CHAIR KNOWLTON: Jolynn, I'm sorry.

MEMBER SUKO: Thank you.

CO-CHAIR KNOWLTON: That's what I say when I can't read them.

(Laughter.)

Go ahead. Got it. Thanks, Jolynn. Sorry.

MEMBER SUKO: So, this is similar to the next measure, sponsored by the American Medical Directors Association. We heard a little bit about this this morning in our introduction.

This is the percentage of the nursing home residents age 65 with persistent indicators of dementia and no diagnosis of dementia on any MDS assessment over the total

of all long-stay residents in the nursing facility who have at least two MDS assessments during the year.

This is a process measure. It is available on electronic clinical data.

In terms of importance to measure and report, as Work Group discussed and as we discussed in our previous measure, dementia is very much underdiagnosed, and prior to diagnosis it increases healthcare costs. So, the Work Group really saw this as important with great potential.

In terms of impact, high and moderate were the ratings.

Let me just look here. In terms of performance gap, we have discussed the performance gap, particularly in the community settings. Again, the Work Group felt like this was pretty significant.

Evidence, this is probably the meat of the discussion and the measure developer has commented post-Work-Group call

this. There are no randomized controlled trials in the long-term care setting. However, there is evidence that the failure to diagnose causes increased healthcare costs. There is evidence that not having a diagnosis of dementia leads to management that is not as effective. The linkage to say that having a diagnosis leads to effective interventions is not as much there.

So, I don't know, David, if we should stop there for comments, discussion.

CO-CHAIR KNOWLTON: Discussion on that point? Any other points? This would be under evidence, right? Importance, of which evidence is important.

Gail?

MEMBER COONEY: The biggest thing that I couldn't find in this was the linkage between making the diagnosis and decreasing healthcare costs, which seemed to be one of their mainstays for why this is important.

CO-CHAIR KNOWLTON: Anybody else?

(No response.)

Okay. Does the developer want to respond here?

MS. VANCE: We feel that is there. There is some evidence in the Singer article as well as evidence in the U.S. Preventive Services Task Force, their systematic evidence review, that shows that at least in the community, again, there are no randomized controlled trials in the nursing home setting. There is the study by Singer that does show that it leads to excessive healthcare costs due to inappropriate care when the diagnosis has not been made. So, we do feel that we have provided that evidence.

CO-CHAIR KNOWLTON: Anybody else on this?

(No response.)

Okay. We can vote on it. Voting on evidence.

(Vote taken.)

I can't ready the number. Are you

still missing some? Missing one?

MS. THEBERGE: We need one more.

All right. Fourteen yes; 8, no, evidence does not meet guidance, and 1, no, insufficient.

CO-CHAIR KNOWLTON: Okay. Going on to impact, please. But we are not voting yet. You present the impact.

MEMBER SUKO: Oh, in terms of impact, the subgroup felt that this had high impact with underdiagnosis of dementia in the community setting, as we discussed in our previous measure.

CO-CHAIR KNOWLTON: Comments at all?

Gail, you have got a comment on it?

MEMBER COONEY: No.

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

(Vote taken.)

MS. THEBERGE: We are at 19, 22.

We need one more vote.

All right. Fourteen high, 9 moderate.

CO-CHAIR KNOWLTON: Okay. We move on to opportunity for improvement.

MEMBER SUKO: And on this, yes, the subgroup that there was significant opportunities for improvement in this diagnosis of dementia.

CO-CHAIR KNOWLTON: Comments?

(No response.)

Okay.

(Vote taken.)

MS. THEBERGE: Twenty-two responses.

All right. Eighteen high, 5 moderate.

CO-CHAIR KNOWLTON: Okay. Reliability?

MEMBER SUKO: So, this measure, it is completely claims-based electronic with precise specifications.

CO-CHAIR KNOWLTON: Anybody on the issue?

(No response.)

Okay, on reliability.

(Vote taken.)

MS. THEBERGE: We have 17 responses, 20. We're at 22.

Nine high, 12 moderate, 1 low, 1 insufficient.

CO-CHAIR KNOWLTON: Validity?

MEMBER SUKO: Face validity was seen as being fairly high. It is hard to manage what you haven't assessed.

CO-CHAIR KNOWLTON: Any comments on validity? Okay.

MEMBER J. BAUTISTA: I have a question.

CO-CHAIR KNOWLTON: Yes, Jocelyn.

MEMBER J. BAUTISTA: I think I read that the specificity of the MDS is about 90 percent, I think I read. So, how do we account for the other 10 percent. So, this

would be 10 percent of patients who score on this MDS, but really aren't the patients that we want to capture. So, how do we account for that?

CO-CHAIR KNOWLTON: It is hard to hear, Jocelyn. Just say it again into the microphone.

MEMBER J. BAUTISTA: All right. So, the MDS has a sensitivity of 90 percent, according to the measure submission. So, there is some 10 percent of patients who will score on this MDS, but who should not have a diagnosis of dementia recorded on the chart. I mean, that is sort of just my simplistic interpretation of that. All right. So, how do we account --

MS. VANCE: I think I can answer that. The purpose of this, the MDS, to explain that a little bit better, it will score something. It is a level of impairment, but it does not give you a diagnosis.

So, the purpose of that is to

bring in a physician that would come in and then they would say, okay, why is this scoring a level of impairment? So, then, the physician would come in and they would do basically differential diagnosis. They would rule out delirium, because you know that is that 10 percent. So, they might have delirium. They might have an infection. They might have some other causes, medical causes, severe depression, something that could lead to that type of scoring.

And then, let's say that they do rule out those other issues or they find that they have those other issues, then that will lead to either with them following DMS-IV criteria to a diagnosis of dementia or not. At that point, then once the diagnosis of dementia would be within the medical record, at that next MDS --

MEMBER J. BAUTISTA: You are basically saying the exclusions account for that remaining 10 percent?

MS. VANCE: Yes, because there could be medical causes for that scoring.

CO-CHAIR KNOWLTON: Ramon?

MEMBER R. BAUTISTA: Just for my education here, how hard is it to give the MDS? Do you need a doctor to do that? Can a nurse do that? And how is it compared to the Mini-Mental Status Exam?

MS. VANCE: A nurse does it, and it is really not that difficult. It is a level of questions that are asked. And then, how the response is, the response is scored and then it is calculated electronically. So, the nurse doesn't have to do like the math. So, it is relatively easy to do, and then the Kappa rating for that is pretty good.

MS. TEIGLAND: I would just say that the nurses are given extensive training on scoring this MDS. It is a science, and there are training classes they take on almost a quarterly basis to make sure that they are scoring it consistently and accurately. And

it just underwent a three-year validation study that was directed by the VA system and RAND corporation. And so, we are really confident that it is a good tool.

The BIMS tool is a validated assessment tool that has been validated against other tools like the MMSE. And so, we are confident that that scoring tool is good.

What you will also see in our measure is that we all want the patient, the resident, to be able to respond. That is how the BIMS is scored. But in cases where patients are too cognitively impaired to actually complete that interview, the nurse, then, does the assessment. So, there are two ways that you can actually be scored for severe cognitive impairment from the resident's perspective as well as from the nursing staff, if the resident can't respond. So, we think we have that covered pretty well.

CO-CHAIR KNOWLTON: Ramon?

MEMBER R. BAUTISTA: We are going

to require nursing home nurses to take the formal training for this and the recertification every "X" number of times. Is that what this measure is going to imply then?

MS. VANCE: No, they are not certified. It is not that difficult of an instrument. CMS runs training courses for Nurse Assessment Coordinators. I mean, this is not given by the LPN. It is given by an RN Assessment Coordinator. Every nursing home has to have one. Or it can be given by the social worker who is also trained. And so, they don't have to get recertified, but they are trained by CMS courses to do so. And then, the Association of Nurse Assessment Coordinators, also, they do have certification courses and do teach it.

CO-CHAIR KNOWLTON: A.M.?

MEMBER BARRETT: I'm sorry, I have to face this way to get to the microphone.

Has the method that you are describing of interview been well-validated to

ensure that there are not healthcare disparities affecting people with communication disorders from deafness, language difficulties, and neurogenic communication disorders?

MS. TEIGLAND: Yes, that was all part of the validation testing for that BIMS tool because, obviously, those are huge issues in nursing home patients, communication issues. Particularly in places like New York City, where I came from, we have multiple languages, and so forth. So, it has been validated. They do require in cases where a language interpreter is required, and so forth, that that is provided. So, that, yes, it is covered well with this tool.

MEMBER BARRETT: I'm sorry, deafness and neurogenic communication disorders?

MS. TEIGLAND: Yes. Yes, absolutely.

CO-CHAIR KNOWLTON: Gwendolyn?

MEMBER BUHR: So, I just wanted to make sure everybody knew that the MDS is being used regardless of the measure and that people are trained for the MDS already. And so, the measure is not going to have anything to do with the MDS being used or not used. It is required to be used by law.

CO-CHAIR KNOWLTON: Dan?

MEMBER LABOVITZ: I am a little concerned about the notion of using this to push for a diagnosis of dementia. Now dementia is a degenerative disease. It means you are declining over time. It is an assessment that can't be performed just once.

This is a measure of cognitive impairment, but it is a measure, I think, even though the BIMS may be very good, I think a staff assessment for cognitive status may not be very good. We may be picking up a lot of patients who have static injuries, old strokes, other things that make them perform poorly on these things, but who are not

demented and where assigning a diagnosis of dementia and improving your performance on this score, on this scale, on this measure, would actually be bad practice.

MS. VANCE: May I address that?

No. 1, we look for two persistent scores on the MDS. So, that means that within 90 days apart having two persistent scores.

Second, like I said, it is an indicator of a level of impairment, but it is not a diagnosis. The diagnosis can only happen by a validated diagnosis by a physician. Nurses are not giving a diagnosis of dementia based on this instrument.

So, it is requiring a physician to come in and do that medically-necessary visit and do a differential diagnosis to come to see if the patient truly does have dementia or what else might be going on that is leading to that scoring of impairment. So, that is the purpose of it. So, that a patient-centric care plan can be developed based on what the

scoring is.

It is unfortunate. Within the MDS, when you have a certain level of scoring, let's say, on the BIMS, there is something that is triggered. It is called the Care Area Assessment, and it will trigger and we will say that there is an indicator that this person has a level of cognitive impairment.

Now you are supposed to address within these Care Area Assessments and say whether you are going to a care plan on that or not. An unfortunate reality is, if the person does not have a diagnosis to go with some of that indicator, the nursing can -- and it is a sad reality, that is why CMS came up with their nursing home measures, which ours were trying to be similar to -- they can say, well, there is no diagnosis of dementia. Therefore, we are not going to create a dementia patient-centered care plan.

And so, one of the major purposes of this measure is then to ensure that we are

raising awareness of this enough that, when a person has this scoring two MDS assessments in a row without a diagnosis of dementia, that you must get a physician in there to look at this person and see why they are scoring the way they are on this BIMS.

So, it is not saying we are pushing that they have a diagnosis of dementia. But if they wind up having dementia, then you want to see it and you want to see a patient-centered care plan around the dementia, the level of dementia they are in, advanced directives, appropriate care, appropriate goals for that person, and leave it that way. And if they have some type of medical issue that is leading to that scoring, you want to see that addressed.

MEMBER LABOVITZ: I see the point in making a diagnosis of dementia and having it done by somebody who is qualified to do it. I just wonder, though, if somebody has a static encephalopathy, they are stable. They

are not demented, but they are cognitively impaired. Does the doctor have to come in and say every time, "No, this patient doesn't have dementia."? When the doctor does come in and say the patient doesn't have dementia, the patient has something else, what happens? The measure still dings the providers here.

There is no exclusion for that. You come along and you say, no, no dementia and, boom, you get dinged next year, too. And you ask the doctor to come back. "Is there dementia?" "No. I told you last year." Well, no you have to do it again.

CO-CHAIR KNOWLTON: Okay. Developers?

MS. TEIGLAND: Yes, I was just going to say that this is another one of those measures that we don't ever expect to be zero. And it is consistent with some of the other CMS measures. One I can think of is depression without antidepressant therapy. Depression is defined by you are having some

symptoms of depression. You are crying. You are tearful. You are sad.

And it is not a definitive diagnosis of depression. These are indicators where you are going to benchmark. You are going to look at your rate compared to other nursing homes with residents like yours and say, "Gee, maybe we are underdiagnosing here."

And the whole point of it is that it triggers a whole different set of reactions by the nursing staff that leads to better care for these patients. And I think we have provided lots of evidence about that. It reduces falls. It reduces functional decline. It helps them better diagnose pain because that is huge. Underdiagnosed pain is a huge, huge issue in this population. It reduces hospitalizations and rehospitalizations because they send her to the hospital.

So, the whole plan of care is different when you properly diagnose. We fully understand that we are going to say,

yes, this person has two indicators of depression based on this BIMS score. They are severely cognitively impaired. The MD might come in and say, "No, they don't have dementia." But, most often, the evidence shows that they do; they will.

And if you look at a list, we have excluded delusions, schizophrenia, bipolar. So, we have really tried to exclude all those confounders, you know, which is really a method of risk-adjusting this measure, but it is not going to be zero.

MEMBER LABOVITZ: I am sorry to hold onto the table, but I see a disconnect between what we are measuring and what the intended outcome is. I completely agree that encouraging nurses and nursing homes and other providers to focus more clearly on the issues related to dementia is important. But I would suggest that what this measure really does is detect cognitive impairment, and it ought to be a cognitive impairment measure. You might

be severely impaired for other reasons than dementia and get no benefit from this as it is constructed. You don't get any of the stuff. This doesn't drive towards that.

And I see the problem, but the measure, by insisting that it lead to a dementia diagnosis, misses out on opportunities and also generates lots of extra work for people who have to be recertified constantly for not having dementia.

MS. TEIGLAND: I think one of the problems is that we have to work with the system that we have within long-term care. And so, with the MDS, with the BIMS, et cetera, we only have scoring for dementia. And so, our system is somewhat limited and not as exclusive as you could get in different settings.

And we know that dementia is a problem. We have numbers of dementia. We have been able to find evidence for numbers of dementia and Alzheimer's disease and defined

evidence for all types of cognitive impairment. It was also more difficult. So, we have to refine our measure to the evidence that we could get and with the systems within long-term care.

So, while I may agree that the perfect measure would include all cognitive impairment, it is not quite possible within the setting that we have and the limitations within our setting possibly to do that.

CO-CHAIR KNOWLTON: Salina?

MEMBER WADDY: I completely agree with Daniel. Those were actually the two points that I brought up on the work call in terms of how accurate is the diagnosis and would it be more beneficial to have something that is less specific.

And I completely understand the points that you are bringing up as well. And so, my major question, I guess, to Christie would be, you say you aren't going to capture 100 percent, but are you closer to 99 or are

you closer to 10 or 50?

MS. TEIGLAND: I think that all we know is what the previous research has shown us and what the U.S. Preventive Task Force found, which is you anywhere from 50 to 70 percent of dementia goes undiagnosed in this population. It is worse in nursing homes than where a lot of these studies have been done.

Let's not forget that the BIMS was just put into the most recent version of MDS, MDS 3.0, because it is a validated measure of cognitive impairment. The measure they were using before was pretty loosey-goosey. It looked at memory, short-term memory, decisionmaking ability. And so, it wasn't as precise.

So, I feel pretty comfortable now that this BIMS score, which has been extensively validated in every setting, is a good measure of cognitive status, but it is not a diagnosis of dementia.

MEMBER WADDY: Right, and that is

the major issue that I am having. But if you have those 50 to 70 percent that are not diagnosed, by implementing this, how much do you all anticipate possibly moving the needle? I mean, I know that you can't really answer that question until it is implemented, but that is --

MS. TEIGLAND: Well, that is what we want to see by implementing the measure and being able to test it. I mean, we feel that, if it is implemented and you have a physician that comes in and is going to rule out medical causes, and the evidence says that there is this huge, huge level of dementia that is going undiagnosed, that we are going to capture a lot of undiagnosed dementia.

And again, that is empirical. So, nobody studied this. Nobody has done it. So, I can't tell you that the evidence leads to this. I am just saying that we have the evidence that shows that you have such a large population of persons with dementia in long-

term care that are not being diagnosed. We know that, by looking at the data, we expect this explosion of Alzheimer's patients, and they are in our setting.

So, we feel that we are going to capture a great deal. But, until the measure passes and we are allowed to start testing it, I can't tell you, which is why I am glad -- I like the fact that it would be a limited measure because, if what we are trying to do doesn't work, then the measure is not worth it. But if we can test it and be able to show what we feel will happen, then we are going to have some terrific outcomes.

MEMBER WADDY: I can give you one example. I had a grant from the Alzheimer's Association, and it was dementia. It was based on the dementia population. And we did look at folks who scored severe cognitive impairment and whether they had a diagnosis of dementia. And so, we ended up using the severe cognitive impairment scores because we

only got about 40 percent of the population with a diagnosis and we added about 20 percent more when we added those severely -- and we went back to the nursing home staffs and had them validate that. They discovered those people mostly really did have dementia. So, that is a little bit anecdotal, but it was a formal grant that I had.

CO-CHAIR KNOWLTON: On validity, Therese, then Mary, then John, then Michael.

Therese?

MEMBER RICHMOND: All right. I do share Daniel's concerns. I won't reiterate that.

I would like a point of clarification. So, I realize that this is based on ICD-9 codes. You are saying only a physician can make this diagnosis. So, a nurse practitioner or nursing -- you have been saying that repeatedly. So, I would like clarification on the specificity of the provider.

MS. VANCE: I probably used the word "physician" because I use that generically. But in our guidelines we use the word "practitioner".

MEMBER RICHMOND: So, it is broader than physician?

MS. VANCE: But it is mostly physician -- we have practitioners as members, but we are mostly a physician-based association. So, I tend to use the word "physician," though we do have, I would have say almost 20 percent of our members are practitioners. And we use the word "practitioners" in all of our guidelines. So, a practitioner can make the diagnosis.

MEMBER RICHMOND: Thanks.

CO-CHAIR KNOWLTON: Mary?

MEMBER VAN DE KAMP: I was going to speak to the fact that we are limited to the MDS within the skilled nursing. I think, Daniel, I agree with you, but what this does is it really takes the lack of specificity of

that tools and drives it to additional assessment.

If you look at what we are measuring, we are measuring a process that drives more than assessment. This process drives change in patient care. So, there is a quality outcome, and it is not just physicians who are engaged when this triggers; it is the rehabilitation staff as well. So, you have speech and language pathologists and occupational therapists who are then engaged, along with the physician.

I think what happens maybe -- I don't know what the percentage, but we need to find out -- is how many are really with dementia and how many are cognitively impaired that would be a result from some other previous stroke, that we then can identify that, once that pool of patients is pulled together, because now the specificity isn't such that you can really determine the best plan of care for those patients.

And it has been an underplanned care, if you will, because it hasn't been to the trigger to pull it out and have specialists review it. So, I think what your concerns are are all of our concerns in the rehabilitation field, but until we can pull them into a group that we can do more physician, nurse practitioner, clinician, therapist evaluation, that lump stays lumped and doesn't really turn into the kinds of best care that we can do.

So, if I look at a process that drives behavior, this process would do that much more than some of the other ones we have looked at in terms of what happens once you pull that group together.

CO-CHAIR KNOWLTON: John?

MEMBER DUDA: So, while I agree with some of Daniel's concerns, to me, they almost seem irrelevant unless you can demonstrate some reason to believe that this assessment with the denominator exclusions

specified would systematic vary from facility to facility. I mean, no facility has zero. But unless there is some reason that some facility logically would have a lot more than another based on their patient population, then I don't think -- you know, we are looking at the exclusion rather than the rule, you know, the exception rather than the rule. Sorry.

CO-CHAIR KNOWLTON: Michael?

MEMBER KAPLITT: So, here is what I am not clear on, and maybe the developer or someone else here can clarify this for me. The denominator is patients who have had at least two -- you said this in response to one of Daniel's questions earlier -- at least two MDS assessments, correct, over a period of time?

So, my question is, where is the evidence to support the validity of this specific measure as it relates to the fact that what you are measuring are those patients

who have actually gotten MDS assessments over a period of time? So, somebody has gone to that effort. The patient has evidence of abnormality on those, and they don't carry the diagnosis. Okay?

So, we are not talking about capturing all these undiagnosed people who have been ignored or who are not be assessed, or whatever. The question is, where is the evidence that in that population of patients that are actually getting this assessment over periods of time and found to be abnormal, that the population that don't actually get the ICD-9 code put in properly, that that is actually going to make a difference or be valid, make a big difference in the care? That is what I am having a hard time understanding.

Maybe I should have raised it earlier under evidence, but since we are talking about the evidence of the validity, I think it is a reasonable time to bring it up,

because I am still not clear on that.

CO-CHAIR KNOWLTON: Hold before you answer, the developer.

Gwen, go ahead.

MEMBER BUHR: Well, I don't know if this would answer it, but everybody in the nursing home gets an MDS at prescribed intervals. So, it is not that a certain population is getting MDS and others are not. Everybody is getting the MDS.

And so, we already know that. And that has been happening since the 1990s. So, everybody has been getting the MDS. And yet, people are not diagnosed with dementia.

And so something, the doctor assessment or the nurse practitioner assessment after the MDS is what has not been happening, I guess. And also, this new MDS has the BIMS where the other one didn't. But it has always had a cognitive assessment in the MDS, and every single patient gets the MDS.

MEMBER KAPLITT: Yes, but before the developer answers, again, it goes to the question of why is this happening, right? So, you say, well, because certain things aren't happening, I guess, right? But, again, where is the evidence that this is actually going to change whatever the problem is? If the evidence is there -- I mean, again, I wasn't one of the primary, you know, I wasn't on this Work Group. So, I may be missing it. But the question is, where is the evidence that this numerator is valid at addressing this issue?

MS. VANCE: Okay. That has a lot to do with the regulatory guidelines. Nursing homes are surveyed by the federal government under state agencies yearly and more often if there has been a complaint. So, if you have an MDS that has a BIMS score that indicates that there is a level of impairment, and you have a diagnosis of dementia, but you don't have a care plan in place for dementia or a patient-centric plan for dealing with that

dementia, that nursing home would be receiving citations, many actually, underneath that -- they are called F-Tags -- for that negligence in care. So, that is one thing. It is not just leading off the ICD-9 coding.

The other thing, as we know, is that with the physician visits every 60 days, and then to 90 days, that unless the nursing staff is calling in the practitioner to come in for a medically-necessary visit, they are not going to know that something is going on with their resident because that is how the nursing home lives and breathes and works.

So, the purpose of this is, okay, yes, sometimes you are going to have someone who doesn't transcribe something accurately. That happens. But, for the most part, because the evidence does show that that documentation is nowhere within the medical record, we know that people are not making that valid diagnosis. We feel that there is more of a chance to capture the missed diagnosis with

this measure than capture that someone did not do accurate transcribing.

I don't know if that answered your question.

MEMBER KAPLITT: But most of your answer related to something that has nothing to do with this measure, which is that a lot of what you said makes a lot of sense. But the numerator is not the number of patients who did not have a care plan attached after they have had abnormalities on the MDS twice. The numerator is the number of patients that don't have the ICD-9 code.

MS. VANCE: Well, no, not an ICD-9 code, but don't have a diagnosis of dementia.

MEMBER KAPLITT: Based on the ICD-9 code, I mean, unless I am misreading this.

MS. TEIGLAND: The ICD-9 code is just one way to get there. There is also a section where --

MS. VANCE: Section (i).

MS. TEIGLAND: -- Section (i) where you can actually check a diagnosis.

But CMS really prescribes how nursing homes sort of operate, and it is really through this tool. If that diagnosis isn't there, it is not going to trigger that evidence-based practice, following that evidence-based practice guideline for dementia. It may trigger doing some things related to the cognitive impairment status, very different from the very much more comprehensive guideline for dementia.

And the sad reality is they just don't follow that evidence-based guideline unless that thing is triggered. So, that is why the care is not optimal for those patients that are underdiagnosed.

CO-CHAIR TIRSCHWELL: This is a little bit of a background question. So, sort of the target problem is the underuse of these evidence-based dementia care plans? And is that more expensive for a nursing home? I am

wondering what the disincentive to the nursing home is to using them. Do they make more money from Medicare for that? Less? It is the same? It doesn't matter?

MS. VANCE: It doesn't matter. It is the fact that they are looking at things like pressure ulcers and falls and urinary incontinence and things that are right in their face. And this is just kind of slipping through the cracks.

CO-CHAIR TIRSCHWELL: So, it is if they have had to MDS assessments over time, then they would have had to have been evaluated by a practitioner on that every 60-day cycle as well, right?

MS. VANCE: Well, the problem is --

CO-CHAIR TIRSCHWELL: So, it is really targeting the bad practitioners, I mean the ones that are not making that diagnosis that you are thinking is there. I mean, they would have to have been seen in that timeframe

for this long stay by a practitioner, right? No? I thought you said it is every 60 days by law.

MS. VANCE: Well, it depends on where they are within that time, every 60 days, and then to every 90 days. And, yes, you are correct.

Unfortunately, if they are coming in and the resident has recently had a fall or there is incontinence to address, there is this and that to address, and there is a limited amount of time, and they kind of know that there is some kind of cognitive impairment, they don't necessarily -- it is not always right on the forefront. I mean, there has got to be some reason why in the community as well as in the nursing home dementia is underdiagnosed.

And what we are trying to do with this measure is make people look at it. I mean, I don't know the reason why. When you look at that United States Preventive Task

Force study, you know, there is some major reason why, you know, it is 50 to 70 percent within the community in the nursing home that people are not diagnosed with dementia. I don't know why, but we want to put it in their face and make people look at it.

CO-CHAIR KNOWLTON: Gwendolyn?

MEMBER BUHR: I think that one problem is that the nursing home does the MDS, whoever is designated in the nursing home. Those results are not front and center for the physicians. The physician comes in to do their visit, and they don't know anything about what the MDS said unless the nursing home makes some effort to tell them. And so, that is a real problem with the MDS and the physician visits, and maybe this measure will help to make that linkage; I don't know.

CO-CHAIR KNOWLTON: We are on scientific acceptability, validity.

Ramon, you have the final point.

MEMBER R. BAUTISTA: So, as a

practical question, though, what would happen to a patient with traumatic brain injury who does not do well on the BIMS score, but is not demented? Where would they fall in all this, though? It is not in your exclusion criteria. Where would TBI patients fall in? They are not being excluded.

MS. VANCE: They would obviously score poorly.

MEMBER R. BAUTISTA: That's right. Where would they fall in here, though?

MS. VANCE: But, then, that would be obviously diagnosed somewhere else. That would probably be --

MEMBER R. BAUTISTA: But they wouldn't be part of your denominator statement then? They would be, but not of your numerator? You would get dinged, though, wouldn't you in a situation like that?

MS. VANCE: Most of the persons with traumatic brain injury, though, we have exclude if the resident is comatose, but we

don't have traumatic brain injury. Most of the residents in nursing homes, though, with traumatic brain injury are under 65.

MEMBER R. BAUTISTA: Well, you could have --

MS. VANCE: But you might have a couple that are over 65.

MEMBER R. BAUTISTA: They become 65 one day, you know.

(Laughter.)

MS. VANCE: Yes, I mean, that is true, but most of them are under 65 because we are actually doing a study with the younger patient in the long-term care setting. But, I mean, if that is a holdup and that is something that you feel that we need to add to the exclusion details, if that's --

MEMBER R. BAUTISTA: I am guessing statement that may be a catchall would be more helpful because there are many exceptions, much more than what you are listing there as an exclusion.

CO-CHAIR KNOWLTON: Can we move on to the vote? You see the criteria. This is validity. Voting is open.

(Vote taken.)

MS. THEBERGE: Nineteen, 21.

Two high, 11 moderate, 9 low, 1 insufficient.

CO-CHAIR KNOWLTON: Okay. Yes, we keep going. It passes.

Who is presenting this? Jolynn?

MEMBER SUKO: So, on to usability, as we discussed, this is derived from electronic sources. The Work Group did discuss -- in general, felt that it was usable, and Salina's point of having a measure of cognitive impairment was brought up under usability as well.

CO-CHAIR KNOWLTON: Yes, Peter?

MEMBER SCHMIDT: So, I am always concerned when I see a measure where the optimal value is not zero or 100 percent from a usability perspective because, how can you

use that for quality improvement if you don't know what the target is? So, there clearly are non-random variations in the issue, your percentage of TBI patients who meet these criteria, but the exclusion we were discussing; those people are not randomly distributed. So, there won't be a random variation of these people who are pushing this measure away from zero across facilities.

CO-CHAIR KNOWLTON: Ramon? Salina?

MEMBER WADDY: So, just to go back briefly to your previous point on who is diagnosing the patient, I mean, I specifically brought up that point on the call and I was told by -- were both of you on the call? I brought up that point, and I was told that it was only going to be physicians at that point. And so, I am a little bit concerned because it just seems like there are small tweaks around the edges that make me nervous about this element. More of a statement than a question.

CO-CHAIR KNOWLTON: Yes. Anybody else on this? Daniel?

MEMBER LABOVITZ: I love to talk to you about this, David. I think this is a squishy measure.

(Laughter.)

And the question, then, comes, is it so compelling that we can tolerate the squishiness? I think that is a judgment call. There is no evidence here. Is this really going to make the difference? Can we put up with the mess that is going to come in some institutions which may have a lot of TBI patients and others which don't? Can we deal with that? Is this going to hurt us or help us?

CO-CHAIR KNOWLTON: Gail, that was an assertive card.

MEMBER COONEY: It was an assertive card. Other than TBI, what makes it squishy, Daniel?

MEMBER LABOVITZ: Anything that

gives you a static encephalopathy, anything that is not dementia that gives you a poor BIMS score makes it squishy. This measure has no capacity for removing those patients from the denominator year after year after year.

CO-CHAIR KNOWLTON: Perhaps a way to go back to answer that question would be to say, what would make it less squishy? And it would be the inclusion of exclusionary criteria such as stroke and any of the encephalopathies that you talk about that would make it less squishy. Just another way to look at that is to just reverse it. That would answer that question.

Mary?

MEMBER VAN DE KAMP: Yes, I wanted to say, back to your things, the diagnosis is physician-driven or nurse-practitioner-driven. There is no soft edges around that. None of us in the practicing fields can -- and to Daniel's squishy comment, you know, I think it is almost the first step, if you will, to get

to the differentiation. I think exclusion would help. But, also, just because the dementia number, there is no dinging for this one, at least from what I can see. They are not going to say you have more patients with dementia in your nursing home because it is not like some of the other measures we looked at where -- wounds is one that is poorly done because you get a facility that has wounds and they didn't grow them, and they get dinged.

Dementia is one that I don't think there is a ding component. I think it is just a better care component. I really think in the practicality of looking at the broader scope of patients in our nursing centers, working to the exclusions which I think are valid but minor really in the population that we are talking about, that I am hesitant to throw out a measure that I think will improve quality down the line for the exclusions that I think would fall out from the further diagnosis by physician and by therapist.

So, I am hesitant. I am sure some of my frustration is that we don't put something out that we don't is based on an MDS which has a lot of validity to it from certain pieces and we don't start to look at additional pieces because it is not perfect yet.

I think one of the ways -- they are going to have 12 months to come back to us to say, "Oh, it didn't work. It didn't show is the right answer. It isn't right." But I am fearful that, if we don't get out in front of this, we don't start defining dementia in this population, it is a really undercared-for diagnosis in our elderly population.

And so, maybe my passion for improved care is overriding my scientific assessment of the measure. But I think there is validity to what they have said in terms of the volume. And there certainly is an importance to improve the patient management with physician and rehab staff involvement.

CO-CHAIR KNOWLTON: Anybody else?

(No response.)

I have one comment on Mary's point. That is, as a former regulator, do not underestimate the ability of a regulator to ding for squishiness.

(Laughter.)

The issue here is this is a facility-level measure. It could easily find its way into inspection criteria.

MEMBER VAN DE KAMP: Would it be dinged, David, for negativity or for patient populations? I don't know which one -- I find it --

CO-CHAIR KNOWLTON: Well, because it is a facility-level measure, a regulator would ding the facility.

MEMBER VAN DE KAMP: For what?

CO-CHAIR KNOWLTON: For having undiagnosed patients where the implication of this is they should be more properly diagnosed.

I am not arguing against your point. I am just saying don't underestimate that capacity, especially, in my view, for a facility-level measure, as a former abuser.

(Laughter.)

MS. TEIGLAND: So, I think that you are right that a high rate on this measure -- or a low rate, because you want this, this is better quality is you don't have a lot of those people, that that might cause a surveyor to come in and look at that resident --

CO-CHAIR KNOWLTON: That is exactly right.

MS. TEIGLAND: -- and see if they were, indeed, misdiagnosed. But, then, if they weren't, if they had the proper documentation in place, which they should have, they can't cite. But that is the whole point of -- and all of the CMS quality measures work like that.

CO-CHAIR KNOWLTON: And that is not a bad outcome.

MS. TEIGLAND: Right.

CO-CHAIR KNOWLTON: But I go back to the point that Daniel pushed back to me. That is, to the extent it is squishy, to the extent that somebody could get zapped for it --

MS. VANCE: But if you look at the majority of the resident population, I mean TBI is not extremely high in long-term care. It does exist. Encephalopathy, I mean, I am sure it exists, but it is not extremely high.

And when you were talking about the risk-versus-benefit ratio that you were asked to consider, I mean, of course, I am one of the developers. But the reason we did this is we live and breathe this stuff every day. We are there in the facilities. We see the patients suffering because they are not getting appropriate care; they are not getting diagnosed. And we just feel that the benefit of this and this measure clearly outweighs any risk of giving it a try.

CO-CHAIR KNOWLTON: Anybody else?

Salina, I'm sorry, I didn't see your card.

MEMBER WADDY: Even though TBI may not be a large segment of the population in nursing homes, certainly stroke is fairly sizable. In aggregate with a bunch of additional diseases, it can be a sizable population.

But I would like to get back to Mary's point because that is actually what is troubling me. This is such a huge problem. It is a huge unmet need. If something isn't done by someone at some point, then it is a lot of patients that are not getting appropriate care.

But the big question is, is this the measure that we should use or is there some recommendation that we can make to make it a stronger or more appropriate measure? I think that is just left to everyone's best judgment.

CO-CHAIR KNOWLTON: Risha?

MEMBER GIDWANI: Yes, it seems to me like we don't want to throw out the baby with the bath water. So, can we just recommend some exclusions and then appropriate, contingent on those exclusions?

DR. BURSTIN: Yes.

MEMBER GIDWANI: Okay.

CO-CHAIR KNOWLTON: A.M.?

MEMBER BARRETT: Just relative to that issue, as a cognitive neurologist, I would remind folks that dementia is a syndrome and not a disease. And so, people can have a stroke and dementia; it doesn't mean that person is not competent to make decisions, can't be static, et cetera.

CO-CHAIR TIRSCHWELL: So, I guess I would suggest that the developers consider adding some fairly, I guess, non-specific exclusion which allows, if a specific other diagnosis is made that can account for the score, that they no longer be counted in the

numerator in future versions of the measure at that particular institution. And that would allow for, yes, everybody to get at least one additional evaluation for the possibility of dementia and, hopefully, more on an ongoing basis. Because even if the stroke patient this year doesn't have dementia, they certainly could have it next year. I mean, I guess if we throw them out permanently, we would lose that possibility as well.

But some additional stipulation whereby, if they have done due diligence and ruled it out, that it no longer counts against them. If that makes people more comfortable, then that might be a way to move forward.

MEMBER WADDY: But how do we move forward? Do we just measure things as is or --

CO-CHAIR TIRSCHWELL: So, this is up for time-dependent --

DR. BURSTIN: No, it is tested.

CO-CHAIR TIRSCHWELL: It is tested

already.

DR. BURSTIN: It is tested.

I guess I have a question for the developers. Is there interest in potentially expanding the exclusions to address this issue? I am not sure I am completely comfortable with the idea of an open-ended exclusion, just because I think that it tends to be pretty imprecise. But I would be curious to hear the developers' response, if that is okay.

MS. TEIGLAND: I think we would certainly be open to adding some exclusions. Our process was that we had an expert panel of geriatricians, who have extensive experience in nursing homes with nursing home patients, come up with this list of exclusions. We thought they were being overly exclusive because they really wanted to limit those residents, those people who end up in the numerator that don't have dementia.

But I think TBI is a good example,

even though the numbers are really tiny, and there certainly may be some other things that they missed. So, I think that is not an issue.

We really haven't tested this measure because that is what we have been throwing out. I mean, we don't know how this would change the numbers of people who are diagnosed. We know there is a big gap, and we hope this would, as all the CMS quality indicators do, cause changes in behavior, which drives better care, better outcomes and better care.

Yes, we have all been hearing about this 30-day readmission rate, right, that they are just implementing? They are dinging nursing homes. But the whole point is that they don't expect that to be zero. They say higher than expected. Everything is benchmarked when we are doing quality measurement. It is all about benchmarking and trying to achieve those higher goals and do

better care and reduce cost, hopefully.

DR. BURSTIN: And just to clarify, the MDS data elements have been validated, which is why the measure is classified as tested, so at least to the moderate level.

MS. TEIGLAND: Right.

DR. BURSTIN: So, you haven't done testing at the measure score level yet. But I just want to clarify, since you contradicted what I said earlier; it is a tested measure.

MS. VANCE: But, as Christie said, we would not have an issue with expanding the exclusion criteria because we honestly didn't think about TBI. We were looking at what large numbers were. But we certainly can add that or add that somewhat statement about, if the physician rules out for a medical cause or a cause, that it doesn't have to be accountable. Maybe we could put doing it yearly or something like that, because a person could get dementia. But we could work with them, a certain type of language that

everybody would be comfortable with.

CO-CHAIR KNOWLTON: So, how do we proceed with that recommendation, Helen?

DR. BURSTIN: It is fine to consider it as part of your voting. It sounds like they are agreeable to add the exclusions; they will work with us.

CO-CHAIR KNOWLTON: Okay.

DR. BURSTIN: And you will get a chance to see those final specs before they go forward.

CO-CHAIR KNOWLTON: Okay. So, in the context of that, can we vote on usability?

MEMBER J. BAUTISTA: So, just to clarify, you mean, if we vote yes, we are assuming they are going to make all those changes?

DR. BURSTIN: Yes, it is contingent on that.

CO-CHAIR KNOWLTON: Okay?

(Vote taken.)

MS. THEBERGE: We need one more.

Six high, 15 moderate, 2 low.

CO-CHAIR KNOWLTON: Okay. Feasibility?

MEMBER SUKO: So, feasibility, these are generated from electronic data sources and, in general, this is the group able to do this, fairly feasible.

CO-CHAIR KNOWLTON: Anybody need to comment on this?

(No response.)

Okay. Let's vote.

(Vote taken.)

MS. THEBERGE: Twenty-one.

All right. Fourteen high, 8 moderate, 1 low.

CO-CHAIR KNOWLTON: Okay. The overall suitability. So, we are at overall suitability for endorsement. Does it meet NQF criteria?

Vote?

(Vote taken.)

MS. THEBERGE: We need one more.

Twenty yes, 3 no.

CO-CHAIR KNOWLTON: Okay. The next is like unto it, and it is Salina presenting on 2092, persistent indicators of dementia without a diagnosis, a short stay.

MEMBER WADDY: So, this measure is very similar, obviously, to the previous measure regarding the underdiagnosis of dementia in patients who have short stay. That is really the major change. It still is a facility measure.

There is a significant amount of data, but, largely, the data wasn't really divided for us between the short-stay versus the long-stay elements. But the group overall thought that there was a significant -- I am trying to find my sheet. The group overall thought that it was an important measure.

CO-CHAIR KNOWLTON: Can I ask a question? It was the same group that considered this? Yes, I am addressing you.

MEMBER WADDY: Yes, it was the

same.

CO-CHAIR KNOWLTON: The same group that considered this. So, it is the same issues?

MEMBER WADDY: So, the comments were pretty -- yes, the exact same.

CO-CHAIR KNOWLTON: Okay. That is what I was trying to find out.

MEMBER WADDY: I didn't think it was necessary to go through it.

CO-CHAIR KNOWLTON: Yes, I agree.

CO-CHAIR TIRSCHWELL: Do the short-stay and the long-stay, then, represent all?

MS. VANCE: It is exactly the same, except for the length of time that you do the MDS assessment. We made ours consistent, harmonized it with the CMS nursing home measures. So, you will see that the CMS nursing home measures are broken up into short-stay and long-stay because their MDS assessments are done with different timing.

CO-CHAIR TIRSCHWELL: I see.

MS. VANCE: And so, to save time, we would agree to do the same exact expansion of exclusion criteria that we agreed to do with the long-stay measure, because everything within this measure is exactly the same except the timing of the MDS assessments.

MEMBER WADDY: Yes, and they convinced us it was necessary to divide those two things out.

CO-CHAIR KNOWLTON: So, without objection, let's just go right through the voting.

Oh, Ramon, I'm sorry.

MEMBER R. BAUTISTA: So, what is short-stay? On this, what is short-stay?

CO-CHAIR KNOWLTON: They are looking it up, Ramon, and they can tell you offline. I think the issue is it is not defined by the measure; it is defined by --

MEMBER R. BAUTISTA: It is not going to impact, though, on the need for this

measure?

MS. VANCE: No, it is defined by CMS. It is a payment issue. They are being paid by Medicare Part A.

MS. TEIGLAND: Yes, it is 100 days. It is you expect to discharge within 100 days. So, yes, these are paid by Medicare as Part A instead of Part B, yes.

CO-CHAIR KNOWLTON: Okay. Can we move on to the voting?

The first will be on evidence, structure, process, and immediate. Vote.

(Vote taken.)

MS. THEBERGE: Seventeen yes; 4, no, evidence does not meet guidance, and 2 insufficient.

CO-CHAIR KNOWLTON: Okay. Impact.

(Vote taken.)

MS. THEBERGE: We need one more response.

Fifteen high, 7 moderate, 1 low.

CO-CHAIR KNOWLTON: And we are on

now -- what are we on, performance gap? Performance gap.

(Vote taken.)

MS. THEBERGE: Eleven high, 12 moderate.

CO-CHAIR KNOWLTON: Yes?

MEMBER WADDY: So, as we go through these, are we also considering the same exception?

CO-CHAIR TIRSCHWELL: The additional --

MEMBER WADDY: Yes, the additional information?

CO-CHAIR TIRSCHWELL: Yes.

MEMBER WADDY: Okay. Great.

CO-CHAIR KNOWLTON: The exclusionary information is you are talking about?

MEMBER WADDY: Yes.

CO-CHAIR TIRSCHWELL: Additional, yes.

CO-CHAIR KNOWLTON: So, what are

we up to? Scientific acceptability, starting with reliability.

(Vote taken.)

MS. THEBERGE: We need one more

Four high, 17 moderate, 2 low.

CO-CHAIR KNOWLTON: Okay. On to validity.

(Vote taken.)

MS. THEBERGE: One more.

Three high, 17 moderate, 3 low.

CO-CHAIR KNOWLTON: Usability.

(Vote taken.)

MS. THEBERGE: Two more.

Eight high, 13 moderate, 2 low.

CO-CHAIR KNOWLTON: Feasibility.

(Vote taken.)

MS. THEBERGE: One more.

Ten high, 13 moderate.

CO-CHAIR KNOWLTON: Overall suitability.

(Vote taken.)

MS. THEBERGE: Twenty yes, 3 no.

CO-CHAIR KNOWLTON: Okay.

MS. JOHNSON: Okay. Great. You guys have done a lot of work, three measures by 12:30. Yay!

(Laughter.)

Who said we might get out early?

CO-CHAIR TIRSCHWELL: I think you jinxed us, Michael.

(Laughter.)

MEMBER KAPLITT: I would like to withdraw my statement from this morning.

(Laughter.)

MS. JOHNSON: Before we break for lunch, I did want to ask very quickly, going back to the measure that just passed, particularly the diagnosis of dementia, do we have any flavor that that would be a disparity-sensitive issue?

I know, A.M., you have already told us that dementia in general is. Can we also say that diagnosis of dementia may also be disparities-related? Again, it is okay to

say no, but you think it is? Okay.

Okay. I might get with you a little bit later and just see if you can point me to a particular source or something. We are doing some background look at some of these things internally. So, that would be super.

Okay. Great.

CO-CHAIR KNOWLTON: Before we do a break, we want to see if the public has any comment.

MS. JOHNSON: Oh, great. Yes.

CO-CHAIR KNOWLTON: Suzanne gets the credit. She tapped my shoulder.

Any members of the public wish to comment?

(No response.)

Anybody on the phone like to comment?

MS. THEBERGE: Operator, can you open the line?

THE OPERATOR: Again, to ask a

question, press *, then the number 1 on your telephone keypad.

(No response.)

At this time, there are no questions.

CO-CHAIR KNOWLTON: Okay. Then, we will be taking a break for lunch.

MS. JOHNSON: Yes. So, since we are running a little bit behind, we are going to try to come back in a half-hour. So, let's plan to start up again at 1:00.

(Whereupon, the above-entitled matter went off the record at 12:26 p.m. and resumed at 12:59 p.m.)

A-F-T-E-R-N-O-O-N S-E-S-S-I-O-N

1:59 p.m.

CO-CHAIR TIRSCHWELL: All right. Sorry about the short lunch, but we are going to jump right back in, so we can try to get done before the debate starts.

Before we start, Michael, we need to give the developer a few minutes, the AAN I guess, to describe their measures.

MS. SWAIN-ENG: Well, good morning, or afternoon actually, since we are in the afternoon.

My name is Rebecca Swain-Eng. I am the Senior Manager of Performance Measurement Implementation at the AAN.

I also have with me today my colleague Gina Gjorvad, who works with me on performance measurement development, as well as Dr. Christopher Bever, who is the lead and the Chair of our Quality Measurement Reporting Subcommittee.

I am just going to give you a very

brief overview. I know we are trying to get back on time here. So, I will keep it short and sweet. I will give Dr. Bever an opportunity to add any additional comments that he may have.

So, just a brief history of the AAN. It was established in 1948 as an international professional association. We currently have more than 25,000 members who are neurologists and neuroscience professionals who are dedicated to providing the highest-quality patient-centered neurological care.

The AAN has a long history of working jointly with the AMA-PCPI on the development of performance measures. We worked with them most recently on the update to the stroke and the stroke rehabilitation measurement set, many of which you reviewed during the Phase I of this Steering Committee project. We have also worked with them on CPAP eMeasures, imaging measures, and dementia

measures, which you will all be reviewing tomorrow.

Additional measures that the Academy has developed include the epilepsy and Parkinson's disease measures that you will be reviewing today, distal symmetric polyneuropathy measures, and ALS measures. We also have measures in process for headache, muscular dystrophy, multiple sclerosis, and so on.

So, the AAN follows the PCPI measure development process. The measures are developed through a cross-specialty, multidisciplinary work group. The measures are publicly vetted during a 30-day public comment period. Once the measures are approved in the peer review, they are then published in the peer-reviewed journal Neurology.

The AAN began developing measures with minimal assistance from the PCPI in 2008. Our Association was actually the first group

to use independent the measure development process with the PCPI. What that means is the PCPI gives us a little bit of staff support, but the whole process is run by our Association staff. They also help us with the vetting of the measures through the PCPI and the measures are actually approved by the PCPI membership and their Board.

The AAN formed the epilepsy and Parkinson's disease measures work groups in 2008 and 2009, respectively. They were developed to fill a gap in the lack of measures that were available for neurological conditions, to focus on epilepsy and Parkinson's disease specifically.

The measures were designed to identify and define quality measures towards managing and improving outcomes for individuals with epilepsy and individuals with Parkinson's disease. The Epilepsy Measure Development Work Group was chaired by Nathan Fountain and Paul Van Ness.

Joining us on the phone today, hopefully, will be one of our work group members, Dr. Gregory Barkley.

The group actually developed eight epilepsy measures, three of which will be reviewed today. These are the three measures that are in the 2012 PQRS program.

The Parkinson's disease measures were co-chaired by William Weiner and Stewart Factor. Hopefully, joining us on the phone today will be Dr. Weiner. He is currently in an emergency. So, we are hoping he will be able to call in with the change in the time today.

The original measurements that had 10 Parkinson's measures, we will be reviewing six of those today, which are in the 2012 PQRS program.

So, there are a lot of additional things that I could say about how we develop our work group, who is involved. It is a multi-specialty group. But I will just leave

that. If you have any questions, I would be happy to answer any additional questions about the work group compensation.

One thing I will mention is that we would ask that the Steering Committee consider the importance of these measures and the significant performance gaps for each measure. Although the evidence that leads the process measures directly to the expected patient outcomes and improvements is somewhat limited, these measures have the potential to significantly benefit individuals with epilepsy or Parkinson's disease. The benefits significantly outweigh the risk. So, we ask that the Steering Committee consider invoking an exception to the evidence for the measures, as appropriate.

As I mentioned, these measures are in the PQRS 2012 program. They are also in a neuro PI program which is designed and approved by the American Board of Medical Specialties to meet the requirements for

performance and practice, maintenance and certification, the Part 4 requirement. They are currently in use in that program, and we have not seen any issues with implementation or usability of these measures in that program.

So, on behalf of the American Academy of Neurology and our epilepsy and Parkinson's disease measure development work groups, we would like to thank you for the opportunity to present these measures.

Dr. Bever, do you have anything else to add?

DR. BEVER: Good afternoon, everybody, and thank you for letting us present.

I guess, as I know you have a Work Group that has already looked through these measures, and they are not based on the highest level of evidence. You might wonder why we didn't just stop working when we discovered that there weren't A-level

recommendations to base our measures on. There really were a couple of reasons at least.

One is that many of the most important aspects of care, based on clinicians' understanding, are not things on which there have been randomized controlled trials and there is A-level evidence. So, oftentimes, we have to make decisions based on lower levels of evidence. So, we think that measures in those areas are important.

The second is an experience that a number of us had in the Department of Veterans' Affairs system back in the 1990s. I think some of you are aware that the VA went through a transformation under Ken Kizer and others in which measurement played a major part. It was credited with both protecting patients from unexpected or unplanned side effects of the transformation and, also, it enabled the VA system to show that in large populations, diabetes, congestive heart

failure, and other areas, that they really did an excellent job and were at least comparable with the private sector.

I was a neurology service chief during that time, and there were no measures for neurologic illness. So, I believe that we took excellent care of our patients with neurologic diseases, but I certainly had no measures to document that. The fact that the planners in the regional offices, the VISNs, which are the VA's Accountable Care Organizations, had no measures for neurologic diseases meant that they really did not neglect neurology at all, but that certainly was not in the forefront of their consideration.

So, I think the American Academy of Neurology together with patient organizations for neurologic illnesses have worked hard to develop measures for neurologic illness, because we think that it is important in the healthcare reform setting to have

measures related to neurologic diseases.

So, thank you.

CO-CHAIR TIRSCHWELL: Thank you.

So, let's go ahead and start with the first Parkinson's disease measure, annual Parkinson's disease diagnosis review, 1973.

Michael?

MEMBER KAPLITT: Okay. So, this is a measure that is designed to capture patients, the percentage or number of patients with a diagnosis of Parkinson's disease in the denominator who have had their Parkinson's disease annually assessed. So, the measure is whether or not people are doing an annual reassessment of the diagnosis and specifically looking at medication use and looking at the presence of any atypical features.

The rationale behind it is that Parkinson's disease is essentially a clinical diagnosis. There are other things that could be used adjunctively, but none of them are considered standard or accepted by the general

community. So, it is still a clinical diagnosis. And therefore, there is a reasonable rate of misdiagnosis in Parkinson's disease. Measures that could improve the diagnosis rate would, presumably, improve care, making sure that patients get the therapies that they need, on the one hand, but, on the other hand, patients who are misdiagnosed don't get therapies that are either ineffective or might actually be harmful to them if they have another type of Parkinsonism or something like that. So, that is the general rationale.

To get to the evidence point, because we said we were going to start with that, the Work Group reviewed this and, then, the subsequent ones after this. You heard a little bit from the developer just now telegraphing their response to some of the issues that were raised on the call.

While we understood, I think, those points, the major concern with the

evidence that seemed to be fairly universal among the Work Group was that there was none with relation to this point. It wasn't that the evidence was just weak. There really was none that specifically relates to this measure.

So, evidence is provided as to the rate of diagnostic inaccuracy in Parkinson's disease, and that, I think, most people do not dispute, that there is a reasonable rate of diagnostic inaccuracy.

And there was some evidence provided as to how better diagnostic accuracy might be useful. The problem is that there was no evidence provided that any of us could find that suggests that annual review improves the rate of diagnostic accuracy. It is true that atypical features that can develop and question the diagnosis may not be readily apparent in the initial diagnosis, and those things could develop over time. But there was no evidence provided that annual re-review

actually changes the diagnostic accuracy rate or would change practice at all. That is what this measure is about.

And so, while many of us, particularly those us who treat Parkinson's patients specifically, are extremely sympathetic to this and the other measures that might help improve the care of these patients, there was no evidence provided on this point. It is not just that, well, you know, there is little evidence and we should -- but we are trying.

There are real concerns because there was no evidence provided, for example, that a general practitioner or a medical doctor or a neurologist who doesn't have much expertise in Parkinson's, there was no evidence that, if they misdiagnose initially, that that would in any way change by an annual reassessment by someone who may not necessarily be as qualified. In fact, as another member of the Work Group

raised/mentioned in the call, the main study was used in support of this measure, the NICE study, which was a study in Great Britain that relied largely on British data that may not be relevant to the U.S., but also did review some U.S. studies, while that talked about inaccuracies in diagnosis, et cetera, that study actually specifically stated that there was no specific evidence regarding what the optimal re-review rate should be, and that patients should generally be referred to a specialist for this purpose, which has nothing to do with this measure.

So, that was the general view of the people on the call. As I recall, I don't think there was a huge amount of disagreement on this point, and this was our major concern.

CO-CHAIR TIRSCHWELL: Great. Does anybody want to comment on this issue, this seemingly lack of evidence, I guess?

John?

MEMBER DUDA: Obviously, you can't

debate that, but I think, getting to this exception thing, there is never going to be any evidence. You know, nobody is ever going to do a study that takes -- well, for this one maybe I guess you could do a study and take half the people.

But the other point, and I don't know if it is the right time to talk about this, but it is kind of a checkbox thing. A doctor says, "Oh, yeah, I reviewed my diagnosis." That doesn't really mean anything other than they have checked off this box.

CO-CHAIR TIRSCHWELL: So, the connection between this and then some improved clinical outcome doesn't seem like it is there?

MEMBER DUDA: Well, that is even a separate issue. A connection between this assessment and whether or not anything was actually done isn't even there, right?

CO-CHAIR TIRSCHWELL: Right. Well, I think speaks to evidence also.

MEMBER KAPLITT: Right. I mean, if it is just a checkbox, but the point here is not that, "Well, what's the big deal?" The point here is that this is a standard that people are going to be held to that we are going to say this is actually a quality-of-care issue. Well, there is no evidence that it is. And so, that is the concern.

And if there is never going to be any data -- first of all, I disagree. I mean, as you suggest, one could do studies on this; it just might take some time. But, you know, that is not our issue here.

I mean, I am very sympathetic to these, I believe. I see patients all the time who are sent to me for surgery who don't have Parkinson's, and I said, "Why are you here?" And so, I am extremely sympathetic to this, but the evidence isn't there. It is not even close. There is nothing.

MEMBER SCHMIDT: I just wanted to say, John, I am currently running, I am in the

third year of a study where part of the study is expert review of diagnosis annually. And so, you can actually get evidence for this and you can get it funded, because I am writing a --

MEMBER DUDA: And extrapolating that to primary care providers and everything --

MEMBER SCHMIDT: No, no, absolutely. No, it is only expert centers and it is only confirming the diagnosis. But there is evidence, and we do get a couple of people -- and this goes to the annual aspect of this -- we get a couple of people sort of in their first four or five years of Parkinson's disease who get rediagnosed, who get a new diagnosis. But if a patient lives with Parkinson's disease for ten years, do you reassess it at nine or ten? You know, there is not going to be any evidence to support that, even in the experts, even for people who are referred late to an expert center.

CO-CHAIR TIRSCHWELL: Daniel?

MEMBER LABOVITZ: This measure would apply to every physician who takes care of the patient and writes down that the patient has Parkinson's disease. So, the urologist is on the hook. The primary care doctor is on the hook. And maybe there is a really good movement specialist taking care of the patient, assessing, adjusting meds. There is no way for these doctors necessarily to have -- they can't say that they did it. Can they attest to the fact that somebody else did it, they think?

I am worried that the patient may be getting exactly what is recommended by the NICE criteria and still generate a ding.

CO-CHAIR TIRSCHWELL: Saline?

MEMBER WADDY: So, back to your point regarding whether or not people would be able to do studies or interested in doing studies, this has come up several times, including the last time we were here. I think

it would be helpful if there was a way that the NQF could inform some of the funding agencies of important gaps that need to be filled, either through NIH, NINDS specifically, or AHRQ or PCORI. And particularly since PCORI is going through their decisionmaking process right now for what to fund, that could be an important opportunity.

But I don't know what your processes are, and certainly this isn't time to expound upon that. But I think it is really important for us, as a federal agency, to get feedback on major gaps that we can potentially provide some answers.

CO-CHAIR TIRSCHWELL: Jack, go ahead.

MEMBER SCARIANO: Yes, on the patients who I see who likely have Parkinson's disease or actually who I diagnose as having Parkinson's disease, it is usually their initial diagnosis. Oftentimes, I have seen

that on those patients within six months or a year they will come back, and they are either totally better or they are doing a whole lot worse. And because I am not their primary care doctor, oftentimes, there is a lapse in between when I see them and actually diagnose them and then when I see them back. The primary care doctors, they just keep on giving them all the same meds. So, I would think that, at least in an early-on diagnosis of patients, that probably seeing them back every year actually would be a good idea.

CO-CHAIR TIRSCHWELL: Okay. Thank you.

If there are no other comments from our Committee, do the developers want to respond to the assertion of a lack of evidence?

DR. BEVER: So, I mean, I agree with the comments that have been made about the evidence base for this. In the evidence-based medicine world, it is not that there are

double-blind, placebo-controlled trials, and then nothing else matters. There are lower levels of evidence for things, including consensus and expert opinion. And so, I don't think those can be totally ignored, although they are certainly lower.

CO-CHAIR TIRSCHWELL: Gail? And then, Peter.

MEMBER COONEY: What I am hearing is more the question of whether an annual review will improve diagnosis. It seems like we should be able to know that.

CO-CHAIR TIRSCHWELL: And that we don't, apparently.

Peter?

MEMBER SCHMIDT: Yes, so some of these things that are addressed in here are not things you would check annually, like responsiveness to levodopa. You give the patient a levodopa challenge and you don't want a year to see whether it worked.

And then, a number of the other

things are at presentation. So, a number of these things are things that I am not sure that saying this should be done annually is -- there is evidence to indicate that annually is not the right frequency or it is not the right time to assess these things.

CO-CHAIR TIRSCHWELL: And I just want to go back to the AAN for one second. Some of these measures are being used in the -- what is it called again? Well, the PQRS, but, no, the AAN maintenance of certification. So, theoretically, there will be a lot more data coming sometime soon that might inform some of these current gaps?

MS. SWAIN-ENG: Yes, we will have more data. Unfortunately, we have had some technical issues where we haven't been able to pull the queries yet. We had a lot of our technical staff, unfortunately, leave in the last six months. So, we will have that availability later this fall to be able to pull more data from that.

We haven't encountered, as I think I mentioned in my introduction, any usability issues. We get feedback, though, from the diplomats that are participating in the program that the physicians that are using this measure and other measures really like the measure, really feel like it is a valuable use of their time. They can see, once we have completed a program, that they have actually improved the care, according to what they are reporting, based upon using this measure. They are reporting a higher level of accordance with this measure at the end of the period.

And we are providing additional resources for them with case control studies, with other articles, to further inform any gaps that they may have in their knowledge. And then, they are coming back and reevaluating this with this measure again, and they are doing very well.

CO-CHAIR TIRSCHWELL: Okay. Thank

you.

Ramon?

MEMBER R. BAUTISTA: I guess the operative word here is "annual". In fact, if you look at the next seven measures, they are all about annual evaluations for something or annual documentations for something.

The difficulty I have is, well, the last epilepsy measure actually listed documentation. But the question I have is, I mean, how do we know that annual things improve care? Is there data that actually even shows that? I mean, nobody is doubting that there are some undiagnosed epilepsy patients or Parkinson's patients there, but how do we know that annual documentation will improve care?

CO-CHAIR TIRSCHWELL: Okay, John, you had your card up there?

MEMBER DUDA: I mean, I think we have addressed that; we don't have any firm evidence for that. These were developed by

the thought leaders in the field, and that is the best we have.

I do want to point out, though, that like this one says at least annually. So, if they do it every three months, they are not going to get dinged for that, and I think some of the others are the same way.

CO-CHAIR TIRSCHWELL: Gail, do you have a comment?

MEMBER COONEY: The AAN is looking at it from a neurologist's point of view. There it is probably useful. I am not sure that it is broadly applicable to non-neurology practitioners.

CO-CHAIR TIRSCHWELL: Okay. Yes, AAN.

MS. SWAIN-ENG: So, I will just address that first question with the annual time period. It is not that it needs to be done in an annual time period. When we are developing a measure, you have to set a time period for the measure. Typically, the time

period, for example, with PQRS is a 12-month period. So, the annually is just saying it has to be done once during that time period. It is not that we can prove annual is better than triannual or quarterly or whatnot. We just have to set a time period when we are developing the measure. It is just a process issue.

CO-CHAIR TIRSCHWELL: Okay. Well, go ahead, Michael.

MEMBER KAPLITT: I just want to make sure that there is no wrong impression left here, because there are two things. I don't want to get sidetracked.

No. 1, while the annual issue may be an issue, that was not the major problem that the Work Group had. Because we agree, I mean, you could always say any measure that has a time period attached to it, you could say, well, is that the right time period versus a month earlier, versus a week longer, versus whatever.

So, I don't want to leave the impression that that is the major crux of the problem we had here. If there was good evidence that at six months or every two years or something made a difference to diagnostic accuracy, I think we would have been more sympathetic to the vagaries of time.

And secondly, to the comment that not everything has to be randomized, controlled trials, again, that was not the issue. It was not that there weren't three randomized, double-blind studies. There was zero evidence presented at all, nor any evidence that any of us could find beyond simply the expert consensus, which is fine for a societal guideline, you know, for a society guideline, but it is not necessarily fine for the NQF standard, based on our understanding of the NQF standard.

So, I don't want to leave the wrong impression that we are arguing over trivialities here.

CO-CHAIR TIRSCHWELL: Thank you.

Mary?

MEMBER VAN DE KAMP: Yes, just being on the Work Group, I think that our challenge was that just by doing it didn't necessarily improve better diagnosis. Again, if the neurologists are doing it, that probably would be valuable. But if you are looking at an overall population, if someone doesn't see it the first time because of what they do or do not know, are they going to see it the second or the third time?

So, it wasn't, again, as we looked at the broader population of assessment; it was that we felt that we didn't improve the skills of the evaluation just by doing it multiple times without evidence that would show why that would change.

CO-CHAIR TIRSCHWELL: Okay. I am going to suggest we go ahead and vote on the evidence at this point. One is yes, and then 2 and 3 are varieties of no.

(Vote taken.)

MS. THEBERGE: Three yes; 8, no, evidence does not meet guidance, and 13, no, insufficient information.

CO-CHAIR TIRSCHWELL: All right. We are moving on to the next measure then because it did not pass on that first evidence criteria.

So, the second Parkinson's disease measure, No. 1982, Parkinson's disease psychiatric disorders or disturbance assessment.

Jane?

MEMBER SULLIVAN: I think there are going to be similarities here with the previous measure. This is a measure that looks at all people with the diagnosis of Parkinson's who at least annually were assessed for the presence of psychiatric disorders or disturbances.

And I think I will echo what the concerns the Work Group had with this one,

which were similar to the prior one, which was that, while conceptually people felt like there was evidence that this is an important issue, that psychiatric disorders are relatively prevalent in this population, the connection between the annual assessment and impact on patient care was not there.

CO-CHAIR TIRSCHWELL: Okay. Thank you, Jane.

So, yes, I think there is a tremendous amount of overlap. Does anybody have any additional comments that are specific to this measure.

Peter?

MEMBER SCHMIDT: So, just a nuance to this one, there is evidence that, for example, depression is difficult to diagnose in a Parkinson's patient. And the measure didn't include diagnosis using validated tools in Parkinson's disease.

CO-CHAIR TIRSCHWELL: Okay. And does the AAN have any additional response that

is specific to this measure, as opposed to the other ones?

DR. BEVER: No.

CO-CHAIR TIRSCHWELL: Thank you for your brevity.

(Laughter.)

Over here, Jane and then John.

MEMBER SULLIVAN: Yes, and I want to just add to what Peter said. There was some concern in the Work Group that, despite the acknowledgment that depression is difficult to diagnose, that was the numerator of the measure. So, it was sort of it was difficult, but, yet, those were the numbers with which this measure was presented.

MEMBER DUDA: So, I guess related to this one, but the last one, too, to my mind, we haven't really discussed whether or not any of these apply to this potential exemption to empirical body of evidence.

CO-CHAIR TIRSCHWELL: Now would be the time to bring it up if you think it is

relevant to use the exemption. Do we want to review that criteria again? Do we have that, Suzanne?

MEMBER DUDA: If it is judged that the potential benefits to patients clearly outweigh the potential harms.

CO-CHAIR TIRSCHWELL: That is not it, though. There is no empirical evidence, expert opinion, and systematically assessed with agreement that the benefits greatly outweigh the potential harms. Pass? Yes, but only if it is judged benefits clearly outweighed harms; otherwise, no.

So, I guess, what are the benefits? I get that there doesn't seem like there could be much harm from this, but I guess I am not seeing any clear information -- I guess I am trying to avoid the word "evidence" -- of benefit. Quite honestly, I don't know how you can establish that benefit outweighs harms without any evidence. So, it seems a little redundant or circular in some

ways.

Okay. All right. So, John, do you want to invoke it?

Gail?

MEMBER COONEY: Well, I mean, you were asking about benefits outweighing harms. It seems that, without assessment of these issues, there can't be treatment of them, and treatment of them would be expected to be beneficial. So, I think that is the link to outcomes.

CO-CHAIR TIRSCHWELL: Okay. Peter?

MEMBER SCHMIDT: So, this seemed to me to be a measure that could be easily fixed, you know, with addition of -- some of the other measures specify instruments. I think that if you kind of address that a little bit, because there is evidence that it is included in the submission that some of these issues are difficult to diagnose in the Parkinsonian patient.

And so, if you just drew from that evidence what are the validated instruments, and included something addressing that in the definition, that could make this a really positive measure. Because I agree with John that diagnosing these things is really important and can really change -- you know, there have been numerous studies, including a paper that I am a coauthor on that is in submission, that have shown that depression is one of the key drivers of quality of life in Parkinsonian patients. So, there absolutely is a benefit from the assessment.

It is just you wouldn't want to pass a measure that kind of said, well, if you examine the patient's effect and you said they seem to be fine, that you have assessed them for depression.

CO-CHAIR TIRSCHWELL: I guess my challenge to you is how is this different qualitatively than the previous measure, which didn't pass? And let me let you respond to

that, Peter, and then Bill.

MEMBER SCHMIDT: So, the different there is that you could actually, in the previous measure, you could put in UK Brain Bank criteria. But some of those things get fairly complicated. With a lot of psychiatric centers, there are validated, short surveys that you can give to a patient that will diagnose these things.

CO-CHAIR TIRSCHWELL: Okay. Bill?

MEMBER BARSAN: I think the problem with both these things and some of the others that we have looked at before is it is really a two-step process. It is not just one step. It is not one thing leads to one thing. It is one thing might lead to another thing, which might lead to another thing.

So, one is, do you assess it? If so, how do you assess it? Do you document that you assess it? And then, there is the assumption that, if you assess it and you document, that, in fact, you do the right

thing. So, there are really two assumptions.

So, somebody could do an assessment for depression and do nothing about it or do something that was inappropriate for it. And you don't have any way of knowing that just by assessing you, in fact, get a better outcome.

CO-CHAIR TIRSCHWELL: Right. And this is another one of those measures that can be achieved through documentation, only a checkbox measure.

John?

MEMBER DUDA: I guess the other difference between this and the last one is that, obviously, a Parkinson's disease patient, when they come to see the doctor for their Parkinson's disease, that is going to be addressed in some shape or form. When you come to a doctor for Parkinson's disease and you don't know that anxiety is a symptom of Parkinson's disease, and your doctor doesn't know that, it is not going to be addressed.

That is why there are some asleep things and these non-motor features of Parkinson's disease assessments are different than just an annual review of the actual diagnosis.

Like Peter was saying, I think there is pretty clear evidence that these things are not diagnosed; they are underdiagnosed. They are undertreated. Improving that is certainly going to improve --

CO-CHAIR TIRSCHWELL: Thank you.

Michael?

MEMBER KAPLITT: Yes, I mean, I agree with that. I do think there is actually harm potentially, and it is an issue. For example, with the last one, the harm issue, right, and why I don't think this was invoked, is that, again, if you have somebody who is not adequately qualified to do this, and then they do this every year and say, "Yeah, I've done it and everything is fine," it leaves the

false impression of quality that is actually not happening. That could actually harm patients because they think that they are doing better than they were before.

That is maybe a little different than a measure that says you should be assessing for this thing. I am not saying that I necessarily feels there is better evidence in support of this measure, for example, but I could see the argument better about the exception for something like this because you are trying to get people who are less qualified to at least think about it to some degree.

Now, having said that, again, there is the potential for harm because people, as you say might misdiagnose it, because now they are being forced to do something that they are not qualified to assess. That is a much vaguer and tougher problem, but it is somewhat different than I think the previous measure in that regard.

CO-CHAIR TIRSCHWELL: Yes. So, I mean, I guess differences and similarities, the potential harm, Michael, you are suggesting is, if somebody uses a validated instrument but is not really an expert or qualified to use it, then they could be put on antidepressant medications, or whatnot, other psychiatric medications that could really be counterproductive in that case.

John?

MEMBER DUDA: So, as with the exceptions we made for the last couple of measures, would it be possible to change the denominator statement to say something like "all patients with a diagnosis of Parkinson's disease examined by a neurologist"? And then, we would get around a lot of these issues we are talking about. Is this a standard of care that we want to apply to neurologists and not other doctors, and is it useful in that capacity?

CO-CHAIR TIRSCHWELL: I guess

maybe we are straying off of the evidence in that respect, John. So, we may need to come back to that.

I need to ask the NQF staff, what is the process for evaluating the exemption to empirical body of evidence? Do we have to vote on that as a group?

DR. BURSTIN: You decide if you want to -- I mean, basically, if somebody calls it out, it is up to you guys to decide if you want to just vote on it, vote on the exception.

CO-CHAIR TIRSCHWELL: So, we should vote on the evidence, and if you want to invoke the exception, you say, yes, there is adequate evidence.

DR. BURSTIN: Yes.

CO-CHAIR TIRSCHWELL: And if you don't want to invoke the exception, you would say one of the "no" responses.

So, if you want to say that there is an exemption to the requirement for

evidence or that you think there is evidence, you would say yes.

DR. BURSTIN: Heidi is going to explain it.

CO-CHAIR TIRSCHWELL: Oh, okay, I got it wrong.

MS. BOSSLEY: Sorry. We spent a lot of time going through these and it is very confusing.

So, if you think the body of evidence as it stands now supports the measure, then you vote yes, which I am generally hearing the answer is no to that.

Then, if the evidence does not meet the guidance, and there is no empirical evidence that exists, that is the one where then we would move you into the exception vote.

CO-CHAIR TIRSCHWELL: Okay.

MS. BOSSLEY: You will do a second vote at that point.

The last one is just there was

nothing provided in the form or in any way to let you evaluate that measure.

So, if you think that you want to invoke an exception, it should be No. 2 that you are going to vote on.

CO-CHAIR TIRSCHWELL: Okay.

MS. BOSSLEY: Does that make sense?

CO-CHAIR TIRSCHWELL: And then, only if a majority votes No. 2 will we move on to the second vote for the exemption? Is that --

MS. BOSSLEY: That is how we did it with the last Committee, yes.

CO-CHAIR TIRSCHWELL: Okay.

Go ahead, Salina.

MEMBER WADDY: So, what percentage of patients who have Parkinson's disease is their Parkinson's disease actually treated by a neurologist?

CO-CHAIR TIRSCHWELL: Can you give us just one number, Peter? Give us your best

guess.

MEMBER SCHMIDT: Yes. So, 40 percent are not seen by a neurologist. Twenty percent are seen by a neurologist once, and 40 percent get their routine treatment by a neurologist at least annually.

CO-CHAIR TIRSCHWELL: Of the ones who are diagnosed?

MEMBER SCHMIDT: Of the ones that are diagnosed, yes, yes, and misdiagnosed.

MEMBER KAPLITT: I just want to clarify the procedural point that was just made, though. Because, based on what you were just saying about the exception rule, to invoke the exception, we have to majority vote No. 2. Then, that means my understanding of the exception, based on that, means that insufficient evidence is not a criteria to invoke the exception. It has to be that there is evidence that just doesn't quite meet the standard, that there is evidence presented --

CO-CHAIR TIRSCHWELL: No, it could

just be expert opinion.

MEMBER KAPLITT: No, because you -- well, right. So, you would have to have a majority of people feeling that there is some evidence to justify that it just doesn't meet the standard, not that there is insufficient evidence, because she is saying it has got to be No. 2. That is what she just said. I just want to make sure we are understanding this right.

MS. BOSSLEY: So, let's look at how this vote would go if you invoked No. 2. So, this is the question that gets asked. If there is no empirical evidence, it is only expert opinion, and you think it was systematically assessed with agreement that the benefits greatly outweigh the harms, then you would vote -- that is what you would be doing if you voted No. 2 on the previous slide. We would go to this vote.

Suzanne, can you, then, go back one?

So, the insufficient information, No. 3, is that, in essence, there is just nothing to support this measure. It is just a flat-out no.

MEMBER KAPLITT: Okay. So, then, I would argue that we may be voting incorrectly, then, on some of these because I don't think there is a measure that we have seen so far that doesn't have some experts saying, "Yeah, this is a reason to do this." Have we ever seen anything with a zero?

CO-CHAIR TIRSCHWELL: Somebody has to bring up the exemption.

MEMBER KAPLITT: No, no, I understand. I am just saying that I think many of us were misunderstanding the distinction between two and three.

CO-CHAIR TIRSCHWELL: We hear that.

And I guess I also don't quite understand why, if two gets a majority in the first vote, why do we have to vote again at

that point?

CO-CHAIR KNOWLTON: Because you might be saying that, but there isn't systematically applied evidence that would allow you -- it doesn't move to the level.

The first vote allows you to say some people think that there is some evidence there, but they are not necessarily saying there is enough evidence systematically applied.

CO-CHAIR TIRSCHWELL: I see. Okay. All right. So, let's go back to the first vote, if we can. This is it right here. I am not even going to try to explain it. I hope you understood it.

(Laughter.)

One, two, or three, let's go ahead and start.

(Vote taken.)

MS. THEBERGE: We need one more. Oh, there we go.

One yes; 18, no, evidence does not

meet guidance, and 5, no, insufficient.

CO-CHAIR TIRSCHWELL: Okay. So, now we do the second vote, and I am going to read this outloud.

"If there is no empirical evidence, only expert opinion, and that opinion was systematically assessed with agreement that the benefits of the measure process" in this case, "to patients greatly outweigh potential harms," we are answering the question, is there an exceptional and compelling reason that the measure should be considered further? One is yes and 2 is no.

David, did you want to say something before we vote? I apologize.

CO-CHAIR KNOWLTON: Yes, I did. I think that this is a new test that we have got to discuss. It seems to me that in this particular case there is a whole bunch of qualifying words in there: "expert opinion," "systematically assessed," "with agreement of benefits, and "Is there an exceptional

underlying and compelling reason?" From where I sit -- and NQF can tell me I am wrong -- but from where I sit, this is meant to be a very high test. I don't think it is being met in this case.

So, I don't want us to just go, I guess because we voted on this No. 2, then I guess this is an "auto in." That is not the way I read this.

DR. BURSTIN: It is not, although somebody has already asked that the exception be invoked. So, you guys can just do a vote on it; that's all. But it is still an exception.

CO-CHAIR KNOWLTON: Right.

DR. BURSTIN: And I think what David said is clear. It is not something we do as a routine course, but when there is compelling evidence that really risks outweigh benefits.

CO-CHAIR TIRSCHWELL: And the exceptional and compelling part, I think that

is an excellent point, David. Thank you for bringing that up.

Somebody -- I can't remember who it was -- referred to the NICE guideline from the UK where maybe some of these things were talked about. I guess I don't know -- I am sure this measure was probably included in that as well.

But anybody have any comments about exceptional and compelling?

Salina, you were first.

MEMBER WADDY: Not on that.

CO-CHAIR TIRSCHWELL: Okay. Peter? And then, Jane.

MEMBER SCHMIDT: So, I think it would be safe to characterize the process that resulted in the paper by Eric Chang as expert opinion being systematically assessed. So, unless it requires us to systematically assess it, I think that this meets that clause.

You know, I agree with John. I think that this is dramatically

underdiagnosed. It is a huge factor in quality of life for people with Parkinson's disease.

You know, if you look at the standardized instrument scores for people who are experiencing psychosis or depression or anxiety, it has a terrible impact on them, worse than increasing motor disability. And so, there really is a compelling reason to assess these, to endorse the assessment of psychiatric disturbances.

CO-CHAIR TIRSCHWELL: Okay. Jane?

MEMBER SULLIVAN: Peter provided the information I was looking for.

CO-CHAIR TIRSCHWELL: Bill? And then, Risha.

MEMBER BARSAN: Yes, I don't know. Again, it is one thing to measure. It is another thing to know that anything good was done by measuring it. And so, there are two -- if there were just one leap I had to make, that would be one thing, but these are two

leaps I have to make, and I just have a hard time making that.

CO-CHAIR TIRSCHWELL: Risha? And then, John.

MEMBER GIDWANI: I have the same concern as Bill. I also have the other concern of whether, given the fact that it was brought up that psychiatric disorders can be difficult to diagnose in Parkinson's patients, whether a neurologist, if we do limit to only neurologists, would have the tools necessary to be able to properly make this assessment or whether it would need to go to a psychiatric professional.

CO-CHAIR TIRSCHWELL: Let alone a primary caregiver, who is theoretically included in this measure as well.

John?

MEMBER DUDA: Remember, Boarded neurologists are boarded in psychiatry and neurology. So, we all have to have some psychiatry training and expertise.

But back to Bill's comment, I mean, I think that this and the other measures may all fail for other reasons. But, as I understand it now, the only thing on the table is whether or not we are deciding that the lack of evidence, you know, systematic evidence that supports this is adequate to deny it, not these other concerns that I have for this measure and all the other measures.

CO-CHAIR TIRSCHWELL: And also that the benefits greatly outweigh potential harms, so another criteria here.

Sorry. Were there any other comments? Risha?

MEMBER GIDWANI: Just a point of clarification. When we say "benefits," do we mean benefits in terms of patient outcomes or in terms of processes of care?

DR. BURSTIN: It is left open, to patients.

CO-CHAIR TIRSCHWELL: Opinion of benefits is my guess.

Daniel?

MEMBER LABOVITZ: I think my willingness to say that there is an exceptional and compelling reason to do this depends very much on who we are asking to do it. If we are asking primary care doctors to be doing this, I think we are going to cause a lot of harm. If we are asking neurologists to do this, and we are talking about 40 percent of the population I guess, because there is not going to be a reassessment after the second diagnosis in the other 20 percent, I am open to that. I would be very interested in hearing further discussion on that point.

But I need to know before I vote on this, can this measure be modified so it is just neurologists?

CO-CHAIR TIRSCHWELL: Okay. Can we throw that one over to the developers?

DR. BURSTIN: No, it is not something we do.

CO-CHAIR TIRSCHWELL: It is not

something we do? What? What is not something we do?

DR. BURSTIN: In general, measures are not to specific specialties. They are at the patient level. They apply to the patient.

CO-CHAIR TIRSCHWELL: They can apply to facilities or clinicians --

DR. BURSTIN: Yes, so clinicians broadly.

CO-CHAIR TIRSCHWELL: -- but not subtypes of physicians?

DR. BURSTIN: Correct.

DR. BEVER: So, would it address the concern if we added to the measure validated instruments that the provider could use?

CO-CHAIR TIRSCHWELL: I am sure that would help, but the NQF is suggesting that we still need to leave it open to all individual providers.

Man, the cards keep going up. David, John, Salina, Peter.

MEMBER HACKNEY: I guess I am a little less concerned, unless I have misunderstood practice patterns, but I see some value in having either a primary care doc or some other physician who is not a neurologist or psychiatrist do the evaluation, and particularly if they have a validated tool to use. And if they think it is abnormal, do they just go ahead and treat or does that spark a referral to someone who is a mental health expert? That might be the appropriate way to go. But if the concern is a PCP may think they have made a diagnosis of depression and treat them with drugs without ever checking, I agree that is an anxiety. I just don't know how many people actually do that.

CO-CHAIR TIRSCHWELL: Well, then that is the second leap of faith that I think Bill has referred to and is worried about.

Who was next? John, did you have another comment?

Salina?

MEMBER WADDY: I mean, that was actually my concern when it was previously mentioned that we limit this to neurologists. I mean, they are, hopefully, more likely to diagnose psychiatric disorders in their Parkinson's patients than the primary care. So, are you really saying that you want to apply a level of quality to the people who are more likely to make the diagnosis.

So, it seems that it will be appropriate, instead, to say clinicians who are seeing Parkinson's patients for their Parkinson's, something along that line, rather than just saying a neurologist or PCP. Does that make sense?

CO-CHAIR TIRSCHWELL: I don't know how you figure out whether they are seeing them for that diagnosis.

MEMBER WADDY: Well, I guess if they are checking off like for the diagnosis code, but what you wouldn't want is -- and that was brought up before -- someone who was

seeing them for a fractured hip and then trying to go through all these permutations that they may not be qualified.

CO-CHAIR TIRSCHWELL: Sure.

MEMBER WADDY: I don't know the wording to tease it apart, but teasing apart those two types of clinicians, ones that are seeing a Parkinson's patient, but not for their Parkinson's.

CO-CHAIR TIRSCHWELL: Helen, can you comment?

DR. BEVER: So, the measure applies only when the provider is billing for Parkinson's.

MEMBER WADDY: That is what I would think.

DR. BURSTIN: It already is, though.

MEMBER WADDY: Okay.

DR. BURSTIN: Yes.

MEMBER WADDY: Okay.

CO-CHAIR TIRSCHWELL: Okay. So,

that is already in place.

Peter? And then, Daniel.

MEMBER SCHMIDT: So, in the UK these assessments are done by geriatricians. You will note that it is assess for psychiatric disorders, not diagnosed with a psychiatric disorder.

I personally think this would be a better measure if you grouped some of these together and said that is an indication to refer somebody to an expert.

But the assessment for psychiatric disorders is routinely by geriatricians in the UK system. That is a very strong evidence-based guideline that they have adopted there.

CO-CHAIR TIRSCHWELL: Okay. Daniel? And then, Gwen.

MEMBER LABOVITZ: It sounds to me like perhaps NQF endorsement is really a very broad brush. It is a broad stroke meant for the population of caregivers, physicians and nurses across the country, regardless of

discipline. It is not set up for this sort of thing.

The American Academy of Neurology has already put this out and is using it, and the doctors who are using it like it. I support that. I think that is terrific. I think it is not only a useful measure for those doctors, but it is also a pedagogical tool.

But if we expand this to an NQF endorsement, then everybody has got to do it. I just don't think the measure is ready for that or appropriate for it.

CO-CHAIR TIRSCHWELL: Gwen? And then, John.

MEMBER BUHR: So, somebody was talking about the diagnosis of depression and whether you would then refer them to a specialist. I think that most commonly not. Primary care physicians would usually treat mood disorders or psychiatric disorders regardless of Parkinson's disease. Whether

that is what they should be doing or not, that is what would happen, because most depression is not treated by psychiatrists or neurologists.

MEMBER DUDA: So, in part in answer to your question, you know, this is only the people who claim to be taking care of a patient for Parkinson's disease. The primary care provider who is taking care of the ingrown toenail isn't going to be assessed for this.

I think, again, you said that it is not ready for that setting. That is not the question on the table. This may fail because it is not reliable and valid, but right now are we saying that there is a compelling reason to ignore the fact that there is no empirical evidence to support this from moving forward to further evaluation, not to approval, right?

CO-CHAIR TIRSCHWELL: Any further comments? Gwen, yes?

MEMBER BUHR: So, my question is to you Parkinson's experts. So somebody seemed to say that it was harmful, it would be harmful. That is my question. Is it harmful if a primary care physician is assessing for psychiatric disorders and treating them? Because you are going to assess for it and then you are going to treat whatever you find. Is that going to be harmful?

CO-CHAIR TIRSCHWELL: Or I guess, theoretically, they could assess for it and not find it inappropriately and not treat it appropriately --

MEMBER BUHR: Right.

CO-CHAIR TIRSCHWELL: -- and that would harm the patient as well.

MEMBER BUHR: So, what are the harm concerns?

CO-CHAIR TIRSCHWELL: John? And then, Peter.

MEMBER DUDA: So, I think missing a diagnosis is not -- I mean, it is harmful to

the patient, but it is not harming a patient. Making a wrong diagnosis and treating them inappropriately could be harmful. But, I mean, are we going to say that primary care providers can't assess psychiatric illness? I mean, that is part of their training, right? And we expect them to be able to do that. I don't think there is any difference because it is a Parkinson's disease patient.

CO-CHAIR TIRSCHWELL: Except that these disorders are notoriously hard to diagnosis in Parkinson's disease. I think we heard that as one of the first lines in this whole thing.

Peter?

MEMBER SCHMIDT: Yes, I agree with what John is saying. There is more harm in not looking than there is in looking.

CO-CHAIR TIRSCHWELL: A.M.?

MEMBER BARRETT: I would just make a little comment that depression in Parkinson's disease I believe is associated

with a higher risk of suicide than it is in other age-matched people.

CO-CHAIR TIRSCHWELL: Okay. Anybody else have any further comments prior to going ahead and voting on this exception?

(No response.)

Okay. John, can you take your card down, please?

(Laughter.)

All right. So, let's go ahead and open the voting.

(Vote taken.)

MS. THEBERGE: We need two more. One more.

Okay. Fourteen yes, 10 no.

CO-CHAIR TIRSCHWELL: All right. So, that means we continue.

So, then, who was doing this measure again?

(Laughter.)

Jane? Impact I think is next, right, 1(a)?

MEMBER SULLIVAN: The Work Group felt that there was evidence of high impact in that the developer provided information that 40 to 50 percent of people with Parkinson's do have psychiatric disorders and 50 percent may develop psychotic symptoms, 30 percent hallucinations in the first five years. And 48 to 80 percent of them may develop dementia. So, the group was comfortable that the impact was demonstrated.

CO-CHAIR TIRSCHWELL: Any comments on the impact?

(No response.)

Let's go ahead and vote then on impact.

(Vote taken.)

MS. THEBERGE: Nineteen high, 4 moderate, 1 low.

CO-CHAIR TIRSCHWELL: Okay. The next criteria is evidence of gap, I believe, 1(b).

MEMBER SULLIVAN: There was data

that the developers presented about the population variance in Parkinson's disease in general, but not specific to psychiatric disease in these patients.

CO-CHAIR TIRSCHWELL: So, there was no evidence that depression or other psychiatric diseases are underdiagnosed or there is evidence for that? I thought I heard people saying there was lots of evidence for that.

MEMBER SULLIVAN: There was evidence that they were difficult to diagnose.

CO-CHAIR TIRSCHWELL: Okay. Peter?

MEMBER SCHMIDT: There is evidence that it was underdiagnosed. I am not sure to the extent that it was actually included in here. But if you go through the references, the references do address the NICE guidelines, one of the references, and they address the underdiagnosis.

CO-CHAIR TIRSCHWELL: Okay. Any

other comments about evidence of a performance gap?

(No response.)

Let's go ahead and vote then.

(Vote taken.)

MS. THEBERGE: Nine high, 12 moderate, 3 low.

CO-CHAIR TIRSCHWELL: Okay. So, then, we are moving on to scientific acceptability. I think first is reliability.

MEMBER SULLIVAN: The comments that have previously been made about specifications for method of assessment, the Work Group talked a lot about that, as well as specifications about which disturbances would be assessed.

CO-CHAIR TIRSCHWELL: And so, the Work Group was comfortable with it as it was?

MEMBER SULLIVAN: The Work Group was a little uncomfortable because there weren't recommendations about a particular assessment tool or modalities for which the

individuals will be assessed.

CO-CHAIR TIRSCHWELL: Okay. John?

MEMBER DUDA: Remind me, but I was under the impression that these new things that have never really been tested were not supposed to be assessing reliability and validity.

DR. BURSTIN: We are only looking, really, at -- because it is not tested -- just 2(a) there, 2(a)(1), precise specifications.

CO-CHAIR TIRSCHWELL: So, I guess this, the lack of tools goes to the specifying how you do the assessment or the lack of specification of how you do the assessment.

Bill?

MEMBER BARSAN: I was wondering if the developers would consider putting in some assessments that should be done, recommended assessments, as opposed to -- I mean, otherwise, this could just be another checkbox where nobody really does anything but says, "Oh, yeah, I checked for it."

DR. BEVER: Yes, so will the NQF allow us to specify? I mean, there are assessment instruments that have been looked at. The Committee did not put them in the actual measure.

DR. BURSTIN: I think the only challenge is it is not just depression. It is depression, psychosis, anxiety, apathy, impulse control. So, you are getting into a whole slew of actually -- and we have already endorsed measures that, for example, use the PHQ-9 for depression or some other promised tools.

I guess the question would be, there are so many; perhaps one option might just be to perhaps insert the words "using a validated tool," rather than necessarily getting into listing them one by one.

DR. BEVER: Right. We would be more comfortable with putting it that way, rather than trying to list all the potential instruments.

CO-CHAIR TIRSCHWELL: Jane, go ahead.

MEMBER SULLIVAN: The discussion that the group had was that in some of the guidelines there were specific tools identified, and members felt that in cases where specific tools were recommended that it might be appropriate to suggest "such as," and then list the tools that have already been vetted by other guidelines.

CO-CHAIR TIRSCHWELL: And I would add that in some of the other measures developers have listed some tools, and they say something like "using tools such as, but not limited to," and then a whole list of possible tools to use.

John?

MEMBER DUDA: Just to clarify, we are kind of throwing clinical acumen out the window and we are saying, if you see a Parkinson's disease patient every year, you have to give them a validated tool for

anxiety, a validated tool for depression, a validated tool for psychosis, a validated tool for impulse control disorders. I am not sure that is really where we want to go, either.

CO-CHAIR TIRSCHWELL: Terry?

MEMBER RICHMOND: That was my point exactly. It sounds like, then, you are saying they need to undergo a full psychiatric assessment, the way this is written. So, I am not clear how that numerator statement would play itself out in specifications. I think that is a concern.

CO-CHAIR TIRSCHWELL: I agree. I think it is very concerning. Of course, the alternative, leaving it as it is, is that the physician saying, "Are you having any psychosis, depression, anxiety, apathy, or impulse control problems?"

(Laughter.)

"No? Okay." Check.

So, I agree. It sort of seems like neither seems very satisfactory, on one

hand, or necessarily feasible on the other hand.

Yes, go ahead, Helen.

DR. BURSTIN: It sounds like most of the discussion we have had today so far has been about depression. And I guess I am confused why the measure has all these other psychiatric conditions. Would that be one approach to potentially hone-in on the areas that are most important?

DR. BEVER: I think in terms of the gap in care, probably depression is the largest in terms of numbers. There are other like impulse control things which are --

CO-CHAIR TIRSCHWELL: Peter, go ahead. Put your microphone on.

MEMBER SCHMIDT: ICD is one of the things that I think is the most impactful to a patient's life. You will have people who will gamble away all their savings. And so, that is important to assess for.

(Laughter.)

Depression, it is the most prevalent, and it has a very high impact because it is so prevalent. Psychosis, again, a terrible quality-of-life problem, but lower prevalence.

CO-CHAIR TIRSCHWELL: Salina? Then, Michael.

MEMBER WADDY: Is there a brief screening tool that combines two or three of these together?

DR. BEVER: I think that was one of the problems, was that there wasn't a brief screening tool. But our committee got in a discussion like this of the various things that happen in Parkinson's, and that is how we ended up with this large number of things in the measure.

CO-CHAIR TIRSCHWELL: Michael, go ahead.

MEMBER KAPLITT: So, my read of this is not that they have to do all of these measures all the time, but they can do any one

of them, right? Because it says "example," and then it gives a list, right? So, they could --

DR. BEVER: It is "or". You're correct, it is "or".

MEMBER KAPLITT: It is "or," right.

So, my concern is actually the opposite, which is that I totally agree with the impulse control issue. However, I can't tell you how many times that I, as a surgeon seeing somebody after 10 years of disease, am the first one to ask them about whether they are having issues with gambling or addictions or sexual things, whatever, because nobody asks them about this stuff with their medicines. So, I agree with that.

The problem in my view with the breadth of this thing is that somebody could ask them every year, "How are you feeling? Are you apathetic a little? Are you okay?" And then, they can check off the apathy box

and that is it. And so, it hasn't achieved the goal.

So, I actually think that specifying the measure down to a specific thing would be a very different thing. My problem is with the breadth of this, that it is just too easy to get credit for having done good care when you haven't done good care.

CO-CHAIR TIRSCHWELL: Back to the check --

MEMBER KAPLITT: Right.

CO-CHAIR TIRSCHWELL: -- easily done by documentation alone.

Any other comments? John, do you have something else to add?

MEMBER DUDA: I mean, I agree, but I think, at least in my mind, the intent of this guideline, and maybe the intent the developers can say, but it was not really to assess in the formal assessment way, but assess, you know, ask, "Are you depressed? Are you gambling too much? Are you anxious?",

or however you want to say it to the patient.

But, then, you are right, you would have to specify that it would have to be "and" for each one of those. I don't know if we are developing --

CO-CHAIR TIRSCHWELL: Daniel, Terry, then Bill.

MEMBER LABOVITZ: I am concerned that we are trying to use NQF validation here for something that is really not meant for it. We are trying to make doctors better. NQF is meant, I think -- and as we go through all these processes, we have had to invoke every exception here to get to this level of conversation.

This measure doesn't fit in. It is not like the others. We need to be using other tools to get doctors to do better on care of patients with Parkinson's disease and depression, anxiety, et cetera.

There is, I think, a desperate crying need, and I suspect that there is a

need for better specialty availability for patients with Parkinson's disease. Maybe just being able to see a neurologist would be a good step or a geriatrician. But you may not even have access to that.

I am not sure that this measurement solves any of those problems. The NQF process isn't really set up to handle a sort of, "Gee, I wish we could do better kind of measure."

CO-CHAIR TIRSCHWELL: Terry?

MEMBER RICHMOND: Yes, I continue to have concerns on the specification, and that on top of the fact that we voted in an exception. The two of them are deeply concerning to me.

Right now, I almost feel like we are trying to redesign the measure as a group process instead of saying, what data do we have and does this meet our criteria? So, just a thought.

CO-CHAIR TIRSCHWELL: Yes. Bill?

And then, Peter.

MEMBER BARSAN: I don't want to beat a dead horse, but, I mean, I feel like we are really pounding very, very hard to get a square peg in a round hole, and it is not working very well.

CO-CHAIR TIRSCHWELL: Peter?

MEMBER SCHMIDT: Yes, so all the thing that we brought up in the evidence point are going to come up again as we go through the future points because they are just as big roadblocks to things like usability and the specification, you know, the use of "or" instead of "and". They are all going to come up as we go on.

CO-CHAIR TIRSCHWELL: Okay. Anybody else have any comments?

(No response.)

Let's go ahead and vote on reliability and -- oh, wait, this is reliability and validity?

DR. BURSTIN: Because it is an

untested measure, and there is no reason to split them.

CO-CHAIR TIRSCHWELL: Well, did we have the conversation on validity?

Jane?

MEMBER SULLIVAN: The only other thing I would add, I think I said before, that there was concern in the Work Group that the inconsistency between the numerator, which is people who have been assessed for this and the difficulties for doing the assessment, that there is data to support that it is difficult to diagnose especially depression in this population.

CO-CHAIR TIRSCHWELL: Okay. That sounds like stuff, as you say, Peter, that we have already discussed to a large degree.

Anybody have any additional comments before we vote on both reliability and validity, because there is no data specifically that we are using?

(No response.)

Okay. Let's go ahead and vote.

(Vote taken.)

MS. THEBERGE: We need three more. One more. Can everyone vote one more time?

All right. Five yes, 19 no.

CO-CHAIR TIRSCHWELL: Okay. I think that means we are done with this measure then.

Okay. Moving on to the next, not terribly dissimilar, Parkinson's disease, Measure 1983, Parkinson's disease cognitive impairment or dysfunction assessment.

Risha, do you want to start us off?

MEMBER GIDWANI: Sure. This is another AAN measure. It is also annual. So, it is all patients with diagnosis of Parkinson's disease who were assessed for cognitive impairment or dysfunction at least annually. The denominator statement is all patients that have been diagnosed with Parkinson's. There are no exclusions to the

denominator.

I can talk a little bit about our assessment of the evidence.

CO-CHAIR TIRSCHWELL: Yes.

MEMBER GIDWANI: So, just a caveat in terms of the numbers that I am presenting to you. It looks like one member of the Work Group voted twice. So, sometimes we have an "N" of five; sometimes we have an "N" of six.

CO-CHAIR TIRSCHWELL: We are not focusing on the Work Group voting numbers.

MEMBER GIDWANI: Okay.

CO-CHAIR TIRSCHWELL: So, just give us the words.

(Laughter.)

MEMBER GIDWANI: Okay. All right. So, the concerns that the Work Group raised were very similar to the ones that we have just heard for the last two measures. And that is that the evidence didn't really address the piece here that we are evaluating, and that is cognitive impairment.

There was also a lack of information about how assessing cognitive impairment would actually result in better patient outcomes. The evidence that was provided by the measure developers was really about depression rather than cognitive impairment.

In terms of the quality of the evidence, there were some randomized controlled trials that the developer cited, but those were actually looking at drugs for treating depression, not, again, for cognitive impairment.

There seemed to be sort of a lot of conflation going on between cognitive dysfunction and impairment and other facets of neurologic impairment associated with Parkinson's disease. So, for example, the measure developers also cited a guideline, and that guideline stated that the Mini-Mental Status Exam and the Cambridge Cognitive Exam should be considered as screening tools for

dementia in patients with Parkinson's disease. That was the evidence that was used. Evidence about this guideline for dementia was used to support their measure about cognitive impairment and dysfunction. So, I think it is really a lot of what we have discussed here earlier, is that there may be some face validity here, but the evidence the Work Group felt wasn't really presented regarding cognitive impairment.

CO-CHAIR TIRSCHWELL: Okay. So, again, very similar issues to the previous measures.

Does anybody have any comments that are particular to this one?

John?

MEMBER DUDA: I think it is a harder argument to make that diagnosing dementia in a Parkinson's disease patient affects their quality of life to the same degree that diagnosing depression or anxiety does.

CO-CHAIR TIRSCHWELL: Okay. Thank you.

Does the developer have anything to add before we vote on the evidence in this case?

DR. BEVER: No, I don't think we have anything.

CO-CHAIR TIRSCHWELL: Okay. Thank you.

Then, well, let's just go ahead and vote. Nobody has invoked anything that shall remain nameless.

(Laughter.)

(Vote taken.)

MS. THEBERGE: Four more.

All right. Three yes; 14, no, evidence does not meet guidance, and 7, no, insufficient evidence.

CO-CHAIR TIRSCHWELL: Okay. Then, I think we are done with this measure, too, as well.

I am sort wondering if we could

skip the lunch break now that is on the agenda -- (Laughter) -- and move through a couple of more maybe before we take our afternoon break. Is everybody okay with that?

Okay. Do you want to take it?

CO-CHAIR KNOWLTON: Yes, Jack, you're up, 1985, Parkinson's disease querying about sleep disturbances.

MEMBER SCARIANO: Yes, as you are looking at this problem, actually, what I do in my practice is that I almost function like a primary care doctor. The patient I see are usually sent in from rural areas and also nurse practitioners. So, the ones I am seeing are usually not diagnosed. So, actually, what I see and the problems that actually happen are almost always seeing people who were just initially diagnosed.

In actually through the actual studies of sleep disorders of patients with Parkinson's diseases, it is really a prevalent problem. If you look in the actual medical

literature, there numerous papers, maybe I would say probably 100 papers worldwide that actually talk about this problem.

The overall problem is, what do you do about it? And then, the other problem is, is the sleep disorder caused by the Parkinson's disease or does the Parkinson's disease cause the actual sleep disorder? In the medical literature, this has been looked at numerous times. They all state that the Parkinson's disease is the cause of the sleep disorder. Studies have shown that, if you have Parkinson's disease, you have a higher incidence of having a sleep disorder. And the most common one is excessive or daytime drowsiness.

And they have also shown that, when you compare Parkinson's patients who have sleep disorders versus people who have other chronic illnesses, say diabetes, that the incidence of having sleep disorders is a whole lot higher in the Parkinson's patients. So,

it is a known problem.

What is it caused by? Well, when they look at it, they have seen that the obstructive sleep apnea is not any higher in the Parkinson's patient than it is in the general population. As you are looking at that, you will say, "Well, it is probably an actual central problem," that it is probably a narcolepsy maybe induced by the Parkinson's disease, or actually who knows?

Studies have actually shown that, that there have been some experimental animal studies that have shown changes in small, little neurotransmitters. And, also, in the Parkinson's surgery group they have seen that some patients who have even had Parkinson's surgery, even though the Parkinson's disease hasn't improved very much in some cases, in the other cases the actual sleep disorder has improved. So, there is evidence all over that it is a major problem.

How do you diagnose this? Well,

there are questionnaires out there that you can do. But I think that the questionnaires are more oriented to the Parkinson's clinics. But it is just basic medicine. I mean, if you ask the patient, if you are a primary care doctor, "Do you snore," does he feel drowsy all day long, you know, just the basic questions that you ask to see if someone has any signs and actual symptoms of having Parkinson's disease, I think that is the easiest way to actually diagnose this.

There has always been an idea -- and I had this, too -- that it is the medications that are actually causing drowsiness. But it is shown in actual numerous studies that it isn't the medication, that it is an actual primary sleep disorder.

CO-CHAIR KNOWLTON: So, what did your group do on evidence? Did they have a recommendation on evidence?

MEMBER SCARIANO: Well, the evidence is that the studies have actually

shown this. There are numerous studies that actually show this.

CO-CHAIR TIRSCHWELL: But, Jack, like the other Parkinson's measures that we have discussed already, I think clearly you are describing lots of evidence associated with an increased risk of these sleep disorders with Parkinson's disease. But I guess the question is, is there any evidence, at least this initial question is, is there evidence looking at this measure, which is asking about sleep disturbances and any evidence that that improves patient outcomes? Or is it that same two-step leap that, if we ask about it, we will identify it; we will refer them to the right person, and then they will get the right treatment?

MEMBER SCARIANO: Yes, well, there is evidence of that. Again, there are numerous articles about that. I think that Dr. Miller is the worldwide leader in this, and she done a study actually worldwide. As

a matter of fact, she actually just finished one in improving the outcomes in Parkinson's patients in like China who have sleep disorders. So, there are numerous studies that actually show this. And I think that it is a valid problem and that it can be assessed.

MEMBER DUDA: Correct me if I am wrong, but I think what you are asking is, is there any evidence that this measure will work? I think you will agree that nobody has ever tested this assessment to see if it will change the diagnosis of sleep problems, just assessing them annually, actually. So, it is like the last one; you don't know it is going to work. There is no evidence to say that it is actually going to work.

MEMBER SCARIANO: It actually doesn't say annually. It says at least annually. So, if you see someone one time and you treat it with medication, and then they come back and say, "Well, he can walk better

and he is not shaking, but he is actually feeling drowsy all the time," you know, is it medication or is it an underlying sleep disorder? And that is where I see it.

CO-CHAIR KNOWLTON: Other comments here on evidence?

Gwen?

MEMBER BUHR: It says in here that it is Grade Level D evidence. So, that is expert opinion.

CO-CHAIR KNOWLTON: Peter?

MEMBER SCHMIDT: So, there is an interesting difference here between the recommendation here and what is in the NICE guidelines. In the NICE guidelines, the statement is that, if the patient complains about sleep disturbance, a detailed history should be taken. That is because the problem isn't so much the diagnosis of a sleep disturbance; it is the differential diagnosis of what sleep disturbance it is. So, I think that that is the major challenge here, that

querying about sleep disturbance is not sufficient.

CO-CHAIR KNOWLTON: Anybody else on evidence?

(No response.)

Let's vote on evidence.

(Vote taken.)

MS. THEBERGE: We need three more. One more.

One yes, 18 no, and 5 no, insufficient.

CO-CHAIR KNOWLTON: The next measure is Mary on Parkinson's disease rehabilitative therapy options, 1988.

MEMBER VAN DE KAMP: I, again, continue the concerns that the group felt around the evidence. I think that rehabilitation, obviously, is a critical component.

There is one concern other than the evidence. It is that the numerator would -- or, I'm sorry -- yes, the exclusions,

actually, would be that any patient with a medical reason, not discussing rehabilitation options with patients or caregivers, where the patient has no known physical disability to Parkinson's disease and patient is unable to respond and no informant is available.

I think that is a large exclusion without taking into account that if rehabilitation is needed, an assessment would be needed to determine if that is true rather than an anecdotal or lack of information. So, I think that exclusion, I feel, has significant issues.

But, like the rest of the measures, the evidence around this I think is that rehabilitation is a value. But a checkbox to say that they were asked about rehabilitative services is not going to change the outcome or the quality.

But, specifically, if we were to have feedback, it was that the exclusions may not actually be the right exclusions for a

true assessment of rehabilitation needs.

CO-CHAIR KNOWLTON: Any new arguments on this one?

Peter?

MEMBER SCHMIDT: I am not sure whether it is a new argument, but I do think that this is one of those things where there is evidence where it front of mind to the clinician results in a higher level of referrals. We have seen that. I have evidence on this that I haven't published yet. But it does make a difference.

And there is ample evidence that rehabilitative therapy makes a difference in patients with Parkinson's disease. So, there is a reasonable causal link.

CO-CHAIR KNOWLTON: But, I mean, can you speak to Mary's comment that the exclusionary problem --

MEMBER SCHMIDT: I totally agree with her about the problems with the exclusions.

CO-CHAIR KNOWLTON: Okay. That is my question.

MEMBER VAN DE KAMP: Yes.

CO-CHAIR KNOWLTON: So, that this measure doesn't do it because of the exclusions.

MEMBER VAN DE KAMP: And I just wanted to support Peter, because, I mean, clearly, the rehabilitative evidence or evidence for rehabilitative care is significant.

I guess the question that we had, as the Committee, one, obviously, the exclusions were of grave concern. But, more importantly, there wasn't evidence in here to show us that that bringing it to the referral or bringing it forward increased referrals to rehab. I think that would be great. I mean, I think that is a great thing. I just don't think we saw it.

CO-CHAIR TIRSCHWELL: I just want to make a comment outloud, maybe for the

developers. It seems like in all of these measures there would seem to be a lot more support if the measure not only included assessment but referral for appropriate care, which would, I guess, increase our confidence that that improved intervention would take place. Of course, it still wouldn't guarantee it, but I think it would get us a lot closer.

So, if this one, for example, were that options for rehabilitation therapy were discussed and were identified and appropriate referral was made -- now I think the hard part is that that is a lot harder to measure, and may be the reason why you are not doing that. But I think there is this conflict between what is the important measure to really drive care and what is hard to measure versus easier to measure with the EHR. So, I would just make that comment.

MEMBER SCHMIDT: I just want to say this is almost one that you could do without any exclusions.

MEMBER SULLIVAN: I was just going to echo what David said. It seems like we looked at stroke measures that said, "Referred for rehabilitation," and then there were exclusions in there, people who, for whatever reason, weren't appropriate. But I think that would capture what really we would try to do to effect care.

MEMBER WADDY: Yes, that was the comment that I was going to make, but a lot more eloquently than I would have made it.

But, to me, it seems like for something that is so clear-cut in terms of making the diagnosis and then referring them for therapy, if we can't manage to put that into a single measurement, I don't know what you would be able to do for practically anything in terms of how we practice, because everything is a two-step. You have to diagnose, and then you have to make a decision. So, how can that really be captured in a single measure effectively and

efficiently? I think that you have described that.

MS. SWAIN-ENG: So, I just wanted to respond to a couple of the comments. I know we had talked about this specific exclusion for this measure during the Work Group conference call.

I just want to reiterate the reason why this exclusion was put in. During our public comment period, we received numerous comments from the public, from different physicians, not only neurologists, saying that they felt that an exclusion was appropriate, because initially we didn't have one for this measure.

Because of the number of patients they see who are so early on in the disease course, they felt like it created an undue burden on these physicians to have to discuss rehabilitative therapy options if it was clear in their professional judgment that this patient did not need that discussed at that

time. It does not mean you cannot discuss it with them. That option is always there. But it helps to reduce that burden on those physicians that didn't feel it was merited for those patients.

And the additional exclusion was patients unable to respond and no informant available. Well, if the patient can't medically have a discussion with the physician, you can't discuss therapy options with that patient. That is just a simple fact.

Additionally, I think one of the additional issues was -- I think maybe that was it. I think that was actually it. That's it.

CO-CHAIR KNOWLTON: Peter?

MEMBER SCHMIDT: So, I know that lots of people don't like to refer, but a lot of the leading experts in Parkinson's disease based on academic medical centers will refer their early-stage patients for an

interdisciplinary assessment at the second visit. They confirm the diagnosis, and then the second visit they do interdisciplinary assessment. I think that is the standard of care adopted at most of the leading centers. So, I am not sure that a community physician not wanting to refer is a great way to do that, but to consider it.

And also, another thing is that difficulty with communication is a symptom of Parkinson's disease. Many of these people can receive information, even if they have trouble engaging in conversation. So, I would look for more than just -- you know, speech pathologists are a key component to a Parkinson's team.

CO-CHAIR KNOWLTON: John?

MEMBER DUDA: So, at the University of Pennsylvania, we have prehabilitation where patients are not debilitated and we send them to the rehab. At the Philadelphia VA Medical Center, I must not

be applying standard of care because we just don't do that. And I think there are a lot of centers that don't have easy access to rehabilitative services, don't refer every PD patient within the first year to prehabilitation.

CO-CHAIR KNOWLTON: Peter, anything else? Anything else, John? You're done?

Okay. Can we vote? This is on evidence.

(Vote taken.)

MS. THEBERGE: Ten yes; 13, no, evidence does not meet guidance, and 1, no, insufficient.

CO-CHAIR KNOWLTON: Which moves us on --

MEMBER KAPLITT: well, no, wait. I hate to do this.

(Laughter.)

But I think it is worthy at least of a two-minute discussion about the exception

rule because I do personally, even though I was pretty harsh on some of the earlier things, I think this is in a different category. I think it is worthy of discussion, particularly since the vote was this close. I think it is worthy of discussion, because I think the risk-to-benefit profile here is very different than assessments of, are you diagnosing things properly or not, or whatever, as opposed to are you having discussions about your therapeutic options. I think the harm issue is very different here and, in my view, much less.

I think that people should be talking about it. So, I think it is worthy of a discussion because there may be a few people in the "no" category who feel it is worthy of an exception that would change the outcome here.

Mary?

MEMBER VAN DE KAMP: Yes, I agree. I think that Michael had done it, and I was

going to do it as well.

I think that, back to Peter's point, if the note gets it to the front and foremost, then I think that, whether it is great, a referral would be a much better option.

But, still, I would like to have the caveat, I am still concerned about the exclusions. So, now I am confused. If I vote for the -- you know, if we say it should be an exception because we believe that the good is greater than the harm, I think the exclusion concerns me around preventing some patients from access. So, I guess I am confused.

MS. THEBERGE: Then we can go to specifications.

MEMBER VAN DE KAMP: So, I am okay. I am okay with that. All right. Thank you. Sorry.

CO-CHAIR KNOWLTON: Other thoughts on the exception?

(No response.)

I am misreading because I am reading this document, 1988, but it looks like -- I was just asking Suzanne -- there is a cut-and-paste error because it is saying no evidence, no evidence, no evidence, and then it is talking about sleep disorders.

So, where are we? Is there a belief -- reminder that when we are making an exception, making an exception says there might not be empirical evidence, but there is expert evidence and that it is very clear.

Did the group feel that in this case it was very clear?

MEMBER VAN DE KAMP: I mean, again, to take the expert component to the rehabilitation advantage, the evidence is high. The risk of not providing that to a patient I think could potentially cause deterioration sooner than might otherwise occur. So, I think there is a harm component, but I am in that field.

MEMBER KAPLITT: I think, yes, the

expert evidence is good. I think the harm is low. And I think the lack of people even understanding the role of physical therapy in Parkinson's disease is a huge problem, particularly given the fact that medications do not treat well many symptoms of Parkinson's, and rehab is one of the few things that can be helpful for a lot of things like balance and walking issues, for example, and other things.

So, I think that the expert evidence is adequate from this in my personal view. But even though I think the evidence doesn't meet the normal standard, I think the benefit combined with the expert evidence and the lack, in my view, of harm in this one compared to some of the others to me does rise to the level of exception, just in my personal view.

CO-CHAIR KNOWLTON: Jordan?

MEMBER EISENSTOCK: I was just going to say I was itching to invoke the

exception, too, even beforehand, just in case.

But I really agree with Michael on this. I think this is a slightly different case than some of the other measures that we have examined recently. Excuse the pun, but it sort of a no-brainer. In the benefit/harm situation, I think that we prevent or minimize the use of dopaminergic medications if we stay one step ahead with the non-pharmacologic treatments like PT and OT. So, oftentimes, I will even try this in my practice if I think the patient can tolerate it and hold off on additional dopaminergic medications.

So, I feel pretty strongly that this is a measure we should try to work a little bit further with.

CO-CHAIR KNOWLTON: Anything else?

(No response.)

Okay. We are voting on the exception. Is there general agreement that the quality, quantity, and consistency of the body of evidence meets the NQF guidance?

I'm sorry. Is there an exceptional and compelling reason that the measure should be considered further, yes or no?

(Vote taken.)

And you are down one, Suzanne. You are down one.

MS. THEBERGE: Okay. So, we still need -- okay, there we go.

Twenty yes, 3 no.

CO-CHAIR KNOWLTON: A compelling argument, Michael.

(Laughter.)

All right. As they say, Mary, you are still alive.

We should be on impact.

MEMBER VAN DE KAMP: Well, I think we addressed that.

CO-CHAIR KNOWLTON: Yes, I think you did, too. I would ask you, however, to also, under impact, address disparities, which is where we have been putting that.

MEMBER VAN DE KAMP: I think that disparities that we discussed were around these exclusions and it broadened the disparities. If they have communication or language barriers, if you are discussing it with the patient, those would have to be addressed as well.

CO-CHAIR KNOWLTON: Okay. Can we vote on impact?

MEMBER WADDY: So, what is your definition again for disparities? Is it just diversity --

CO-CHAIR KNOWLTON: Can we hold it, Salina, because I put it in the wrong place.

MEMBER WADDY: Okay.

CO-CHAIR KNOWLTON: We will discuss disparities in the next round.

(Vote taken.)

MS. THEBERGE: I need three more. Oh, there we go.

Eighteen high, 5 moderate.

CO-CHAIR KNOWLTON: Okay. Go back to the performance gap, Mary.

MEMBER VAN DE KAMP: Yes. Again, I think it speaks to the conversation of bringing it to the forefront with the physicians will, then, improve the access to rehabilitation services, hopefully sooner than later, and certainly ongoing.

CO-CHAIR KNOWLTON: And Salina's point on the disparities, did you hear it?

MEMBER VAN DE KAMP: Yes, and I think I was addressing it as well earlier. But I think this is a measure with exclusions that concern me. It is that, if there is a disparity around a language barrier or apparently not understanding, or I understand to some degree what the response was, but it concerns me that we are making a determination of whether a patient or their family member, or a patient specifically can understand before an assessment of whether they have comprehension and the skills to make that

determination. I mean, it is like crossing them off before we assess, I guess.

CO-CHAIR KNOWLTON: Go ahead.

MS. SWAIN-ENG: I can speak to that just very briefly. This is a medical exception. A language barrier --

CO-CHAIR KNOWLTON: Hold on for a minute. Just hold for a second and let Salina respond to the question, so we get the back-and-forth.

MEMBER WADDY: Yes. So, I just wanted to be clear in terms of how you are all defining disparities, at least across the federal agencies there are three components, both minority as well as role versus urban and socioeconomic. And certainly, if you don't have funds to be able to pay for rehabilitative services, then that is a disparity in and of itself. As well, in rural and remote places in the middle of Alaska, seriously, you are not going to find good rehabilitative services. And there are

various reasons why people may not.

And so, I just wanted to know what NQF's definition --

DR. BURSTIN: Much of the work that we have done to define disparity sensitivity was done on race, ethnicity, and language. I think the idea of prevalence and a performance gap and an opportunity for improvement are things that I think would work well across any of those other entities.

And we really just want to get a sense of, really, essentially, is this a measure that should be stratified, so you don't miss out on populations particularly at risk?

CO-CHAIR KNOWLTON: Rebecca, you had a point?

MS. SWAIN-ENG: Sorry. I am just trying to say that, if somebody did have a language barrier and that was an issue, that is not included in this measure. This is a medical reason. So, there is a medical

condition, problem. Perhaps somebody was late-stage dementia with Parkinson's disease and didn't have somebody there with them to either act as their caregiver or they couldn't cognitively because of a medical respond or participate in any meaningful discussion. So, language barrier wouldn't fall underneath this issue. You would get an interpreter, and that wasn't covered or intended by this exclusion.

CO-CHAIR KNOWLTON: Anything else?

(No response.)

Okay. We are on the performance gap, yes.

(Vote taken.)

We still need some votes.

MS. THEBERGE: I need four more votes.

All right. Nine high, 12 moderate, 2 low.

CO-CHAIR KNOWLTON: Okay. So, we are moving on to scientific acceptability, starting with reliability.

MEMBER VAN DE KAMP: This is the area that I think I brought up too soon, obviously, is the exclusion concerns that I have already addressed.

CO-CHAIR KNOWLTON: Jane?

MEMBER SULLIVAN: This is an area that I feel like I would like to say that there is some burgeoning evidence of the neuroprotective effect of exercise. So, in addition to what has already been said, I think that the exclusion of non-motor symptoms, the fact that this is a progressive disease, is compelling reason to look seriously at removing that exclusion.

CO-CHAIR KNOWLTON: Other thoughts on this? This is reliability and validity combined in this particular measure. Anything else?

Michael?

MEMBER KAPLITT: From the developer, whether they are willing to do this or not, because that, I think, is going to

affect a lot of votes.

CO-CHAIR KNOWLTON: Okay.

DR. BEVER: So, what is the specific request?

MEMBER SULLIVAN: The request is to consider the exclusion of non-motor symptoms, the patient who is not presenting with a motor symptom as an exclusion, because it is currently stated that it is an exclusion and there is some concern that has been expressed that this exclusion would prevent somebody from being counseled about rehabilitation until or unless they had some frank presentation of the disease. I think prehabilitation was the term that you were using, would eliminate care for people before they maybe a year down the road were showing frank motor symptoms.

DR. BEVER: So, you think the measure will be used as a guideline, basically, to tell you when you have to do something? And so, the fact that the measure

doesn't -- you are saying the measure would lead somebody not to do rehabilitation in someone with non -- I mean, the exception wasn't meant to exclude that. The exception was only meant, as a quality issue, you are not required to counsel that person. You are saying, as a quality issue, those patients should be counseled.

MEMBER SULLIVAN: I guess my point was that somebody who is sensitized to the disease and the opportunities would probably do it anyway, but the primary care physician who is maybe not seeing a lot of these patients would say, "Oh, well, they are not showing motor symptoms. So, I don't need to discuss rehabilitation with them."

And I would like to advocate that, if they have that diagnosis, even if they are not showing symptoms, they may, and intervening early would have some benefit. So, to take the exclusion off the table.

DR. BEVER: Well, I mean, there is

some evidence, as you point out. I don't think that that is a standard of practice yet, would be my understanding. I don't know; maybe others who deal with Parkinson's patients would want to comment on that.

CO-CHAIR KNOWLTON: I don't understand. Well, let's take a few more comments and then we will come back to you because there is going to have to be some clarity on what the developer is willing to do to meet the exclusion issue that a number of people voting have a concern about.

Peter?

MEMBER SCHMIDT: So, although I know that lots of people like the prehabilitation model, there really isn't evidence for it. And so, the exclusions, as they stand with some nuances around difficulty with communication, I think these exclusions as they stand are in line with the evidence for health interventions.

CO-CHAIR KNOWLTON: Mary, John,

then Dan.

MEMBER VAN DE KAMP: I would just say that discussing rehabilitative options doesn't mean today. And so, I think maybe that is where I am interpreting it. It goes back to, do you refer or do you just discuss? So, I think if you look at it as an opportunity to discuss the possibility of rehab might help you at a certain point, it may be of value. I think it is not saying that you must have rehabilitative services to get a quality check.

I worry a little bit in this, in the exclusions, that you are leaving a lot to a prejudice maybe of the assessor on whether rehab is valuable or not generally and leaving that more to taking it to the evidence around rehabilitation over the course of care within a rehabilitation process.

I don't know; I hear what you are saying about not wanting to have everyone get rehab. You know, we don't need to have a lot

of evaluations for rehab that aren't appropriate because that raises the cost and is of no value.

But this doesn't say evaluations; this says rehabilitation, right? I mean, we are talking about that. So, I don't know. I am vacillating a little, I guess, on that one.

CO-CHAIR KNOWLTON: John?

MEMBER DUDA: So, in my practice, like I said, I don't talk to people about PT/OT and speech therapy if they come in with a benign resting tremor and one extremity that is non-disabling. I don't see the point of that.

In every one of those patients, I do talk about physical activity or exercise. I think we are kind of blurring the distinction here, that a lot of the evidence for neuroprotection and everything is really for an active lifestyle, not for going to physical therapy and getting treatment. There is no evidence that I am aware of that

suggests that this has any effect on the progression of the illness. It affects the functional capacity and things. But if the physical therapist can convince you to do your exercise, sure, but that is not what we are talking about here, right? We are talking about a specific regimen of rehabilitation for a specific deficit. And a lot of PD patients don't have any deficits early on.

CO-CHAIR KNOWLTON: Daniel?

MEMBER LABOVITZ: John spoke my point. I think anybody can advocate exercise. You don't have to get it from a physical therapist.

MEMBER BARRETT: I would actually say that I think that your recommendation constitutes a discussion of rehabilitation options appropriate for that patient's stage of care, and fits a standard of care within neuro-rehabilitation for those kinds of patients.

I would say that you are doing a

rehabilitative option discussion when you do that. When you discuss physical fitness, I would say that that fits a standard of care within neuro-rehabilitation for those patients in general.

MEMBER WADDY: I would actually say you just made this a lot more difficult because separating those two out, whether or not it is just increase in exercise or some exercise regimen as opposed to rehabilitative therapy -- and those are two separate things -- and how this issue is actually addressed by practitioners, does it really reach the level of putting in rehabilitative therapy?

CO-CHAIR KNOWLTON: Michael?

MEMBER KAPLITT: I would also say that the exclusion has a documentation requirement to it. I think that that, to me, is important in giving me a comfort level with this. Because it is like, if I don't give antibiotics before a surgery, a PQRS measure requires me to document it. So, I can't just

choose not to do it and say, "Well, it wasn't important."

So, I think here the documentation requirement is going to be put a little burden on people who come in with just a tremor and every time you have got to say, "I didn't discuss rehab with them because there is no need," but it does at least require people to have documented that they thought about it and why. And so, that gives me a little bit more comfort level.

CO-CHAIR KNOWLTON: Peter?

MEMBER SCHMIDT: So, this is genuinely an area of clinical controversy. This isn't something that we can decide here ourselves.

There is a study ongoing in Australia where they are randomizing people into a group where it is neurologist-directed care versus a team assessment. I wrote the check for that study. So, I have seen everything about it.

It is established clinical controversy. People don't have an assessment. So, it is not appropriate to remove the exclusion and define as quality care to address this at presentation because there isn't the evidence for it.

You know, we may like that idea, and we funded that, my Foundation funded that project because we like the idea of doing this assessment. And maybe in a year this study will be published and we will have one RCT to address this issue with. But today we can't do it.

CO-CHAIR KNOWLTON: So, to the developer -- oh, I'm sorry. Gwen?

MEMBER BUHR: So, now you made me have a question.

(Laughter.)

Thinking about reliability and the exclusions, so if you have convinced me that we should keep the exclusions, are we going to be able to always get the same patients with

these exclusions? It seems like they can be interpreted sort of however you want to because it just says, "example". So, you could just say it wasn't appropriate for that patient, and anybody can have a different reason for why it is not appropriate. That doesn't seem very reliable.

There is a medical reason for not discussing rehabilitation therapy options with the patient or a caregiver, as appropriate. So, you can think of whatever medical reason you want to.

MEMBER KAPLITT: But that is true for a lot of these types of measures. I mean, we are giving people some element of clinical judgment. And that is why I think the documentation requirement at least forces you to give that reason. It is possible over time that that would change. I mean, we have already accepted the idea that we are making an exception and that the evidence is not there, but that we feel it is important

enough.

I think that to mandate this overall for everybody, say that you actually have to have this discussion, you know, it is probably not a big deal. But if you have people who are mute and they show up from their nursing home with somebody from the ambulette service and they don't have a family member there, and you still have to have that discussion, you know, I think that there are enough reasons that, as long as it has got to be documented by somebody, you know, yes, over time that may change, but I don't know that that is a huge harm. I mean, you are right, people could do that, but you could say that about almost any PQRS-type thing. I know that is not exactly what this is, but --

CO-CHAIR KNOWLTON: Other folks?

(No response.)

Developer?

DR. BEVER: So, do you want to vote on it as it is?

CO-CHAIR KNOWLTON: Well, yes, I don't know. What have we got here? What is on the table?

MEMBER WADDY: Well, can I just say really quickly, I mean, there is the exception issue, but I still have an issue regarding the wording of rehabilitative therapy as opposed to potentially exercise. Does it specifically need to be within PT/OT or speech?

CO-CHAIR KNOWLTON: Unless somebody is going to say that they want a specific exception here, I am going to leave it as it is. So, if you want the exception, speak up.

(No response.)

Okay. Then, we are voting on this as is, on reliability and validity, and you are going to take your best shot.

(Vote taken.)

MS. THEBERGE: I need one more.

Eleven yes, 13 no.

CO-CHAIR KNOWLTON: So, we are done with this measure.

Peter, you are up.

MEMBER SCHMIDT: This measure is Parkinson's disease medical and surgical options reviewed, although in the definition it also talks about non-pharmacological treatment, pharmacological treatment and surgical treatment, reviewed at least annually.

So, these things need to be reviewed for the patients who are seen. My first reaction to this was, if the patient is coming to the clinic and you are not reviewing their medical and therapeutic options, then what are you doing?

(Laughter.)

So, there is no evidence for this. It is at least annually. However, I think most patients, we do surveys of centers, and most people will see their average patient every three to four, maybe six months.

So, this is really not very well supported. There is no real evidence around this because you would never get it past an IRB to test not reviewing medical options when the patient comes to the clinic.

CO-CHAIR KNOWLTON: Don't hold back, Peter.

(Laughter.)

MEMBER SCHMIDT: Okay. Could I just say I appreciate that AAN submitted these guidelines, and I think it is very important, but I think this particular one being defined as a quality standard is challenging.

CO-CHAIR TIRSCHWELL: I am just looking at this, and I guess although certainly the pharmacologic especially, but I am guessing that there is a good number that don't have non-pharmacologic or surgical options discussed on an annual basis --

MEMBER SCHMIDT: It is "or".

CO-CHAIR TIRSCHWELL: Yes, and I agree; maybe that is just a suggestion that

needs to go back to the developer, that maybe certain aspects of this are more relevant for a quality measure than others. I don't know.

MEMBER KAPLITT: Yes. No, I think the "or" is the big issue because, if people say, "Yes, your medicine seems to be working just fine," and that's it, that is the discussion, or even if it is not, and then that is it; they satisfy the criteria. So, it is the "or" that is the issue, I think.

And I would argue, just for the sake of maybe brevity or expediting this, that I know that we had said we were going to do evidence first. But the issue which is raised, which I think is an important one, is really more of a performance gap issue, meaning is there really evidence that there is a gap in the fact that, when patients come to their doctor to be treated for Parkinson's, they are not discussing treatment for Parkinson's, right? Is there real evidence of a gap?

So, I mean, I would propose that maybe that be the first thing we discuss because that was the issue that was raised.

MEMBER SCHMIDT: So, in fact, there is evidence that it is being discussed because you can see patients having escalating doses and adjunctive therapies added on in the community setting. And quite often, when patients are referred to expert neurologists, their medications are reduced, not increased, which indicates that somebody is thinking about their medications, just getting it wrong.

CO-CHAIR KNOWLTON: Anything else here?

(No response.)

Let's stay on the evidence first. The gap issue we could discuss, if you want to, but -- the rule of the Chair, we will go with evidence first. Let's vote on evidence. Voting on evidence.

(Vote taken.)

MS. THEBERGE: Two more responses.

Two yes; 16, no, evidence does not meet guidance, and 6, no, insufficient evidence submitted.

CO-CHAIR KNOWLTON: Okay.

CO-CHAIR TIRSCHWELL: So, we are back on time.

(Laughter.)

Can we take a 15-minute break? Is that okay?

CO-CHAIR KNOWLTON: Sure.

CO-CHAIR TIRSCHWELL: Yes, go ahead.

MEMBER WADDY: I just think that, with the previous one on rehabilitative services, to the developers, I think it is really unfortunate what happened with that one. I think it is really important that, if there was a way to somehow address the criticisms that you heard, I mean, I think it would be really of value to revisit in some subsequent time period.

CO-CHAIR KNOWLTON: Well, I am glad you said that, Salina. I would add to that, I am probably the least clinically knowledgeable here, but I think the developer, just looking from this side of the table, people are all saying that these are important things and requiring attention. This is the first real shot at trying to pay attention to them in a structured way. And the NQF standard is a high standard.

But I don't hear any of these things where people say, "Now why are we even bothering with this?" People were very supportive. It just didn't quite meet the test. So, I hope these will be things that we will continue to work on.

DR. BEVER: Yes, I think the challenge at the developer level is that there are different criteria at each level that we are working on these, and each group has their own thoughts about how they should be crafted and different considerations. And so,

navigating that has been challenging.

CO-CHAIR KNOWLTON: One of my comments to Suzanne during this debate was democracy is messy. You know, getting consensus through this process is a very high bar. But, at the end of the day, hopefully, it gets better.

I have seen these debates before. I have been on a number of these. This was a good one and a rich one, and it was a very positive one for these measures. I have watched enough measures go down in flames; that isn't what happened here. So, there is a lot of support for these measures. So, I hope you won't be disheartened. That is just an editorial comment from me. Don't be disheartened. These are good measures. They need some tweaking to get through the consensus process.

DR. BEVER: Thank you.

MEMBER WADDY: I agree with that. I mean, I just really think that that is such

an important one, in particular, that if there was a way to take some of our comments and tweak it, because this isn't a measure where you just have to throw it out and start completely over. I think that tweaks, small tweaks, can really change how the measure is viewed.

MS. JOHNSON: And just to remind everybody, like we did last time, tomorrow we will have a little bit of time for you guys to weigh-in in terms of ideas for future measure development. So, that might be something, and we always write those up and put those in our reports. So, we would encourage the developers to take a peek at that as well.

(Whereupon, the above-entitled matter went off the record at 3:02 p.m. and resumed at 3:20 p.m.)

MEMBER RICHMOND: Okay, let's go ahead and get started again.

Raj Sheth will be presenting the next, our third-to-last measure for the day,

1814. We are switching to epilepsy. This is counseling for women of childbearing potential with epilepsy.

Do we need to introduce the developers? It is the same developer for all three of these.

Did you have any different comments about the epilepsy measures as opposed to the Parkinson's ones.

DR. BEVER: No

CO-CHAIR TIRSCHWELL: Okay. Great.

So, Raj, go ahead and start us off with an overview and then right into the evidence.

MEMBER SHETH: Thank you.

This is a pretty big issue, and it is not one where you run into the typical challenges of is or is the patient not depressed. I mean, the numerator should be relatively easy, at least from a pregnancy perspective. You are either pregnant or not,

obviously. And so, from that regard, it is very important.

There are probably half a million women with epilepsy. The amount of controversy that exists about pregnancy, contraception, breastfeeding is huge.

There are also two people in this. There is the fetus and the mom. So, it really has a big impact. Teratogenic effects on the fetus really have long-term consequences and huge costs. These are going to be 30-, 40-, 50-, 69-year expenses. So, the overall impact is quite significant.

And the rationale that is provided by the developers is that the performance gap is that only 2 to 20 percent, between 2 and 20 percent of women --

CO-CHAIR TIRSCHWELL: We are not on performance gap. It is this --

MEMBER SHETH: So, that is the overall introduction.

CO-CHAIR TIRSCHWELL: So, I guess

we are looking for evidence that this measure, as crafted here, will have a positive effect on patient outcomes.

MEMBER SHETH: And the evidence is, like the discussions that went before, it is very scant and not really highly --

CO-CHAIR TIRSCHWELL: It is not a direct link. It is through a couple of intermediate assumed processes?

MEMBER SHETH: That is correct.

CO-CHAIR TIRSCHWELL: Okay. Thank you, Raj.

Anybody else want to comment? Dr. Barsan?

MEMBER BARSAN: Yes, I think the issue is that nobody doubts that this is important and nobody doubts that this is a critical issue. It is a question of, does this, as outlined, doing these things, is that going to really make a difference? Is that going to affect an outcome at all?

And so, trying to determine the

assessment with the outcome. I mean, the same problem with the other assessments. And so, I think that is really where the issue comes in.

The other thing that we talked about, too, is, is it sufficient to say, "Here's a website."? Is it sufficient to give handout materials? Does it have to be a half-hour discussion? I mean, you know, there is not a lot of discussion about what is adequate in terms of that. So, that is part of the issue, too.

CO-CHAIR TIRSCHWELL: Daniel?

MEMBER LABOVITZ: I am just warning you that I am going to invoke an exception on this one.

(Laughter.)

I think it is really important, and I don't need much evidence to believe that talking to patients about this issue makes a difference, both in terms of patient behavior and in terms of knowledge and the provider's

awareness.

I have screwed this up. It didn't lead to any disaster, but I failed to have the conversation. I feel bad about it. I think it is a quality measure that needs careful scrutiny.

CO-CHAIR TIRSCHWELL: Michael?

MEMBER KAPLITT: I mean, before we get to the exception point, I think that there is some evidence presented here. I mean, again, it is not randomized controlled evidence, but there is evidence here compared to some of the other measures we talked about that are directly on point. I mean, there are several surveys, for example, that they cite, large surveys, of women of childbearing age who report that they feel, you know, a large percentage feel that they are not being adequately informed. Well, I guess that is more of a performance-gap issue.

CO-CHAIR TIRSCHWELL: Yes.

MEMBER KAPLITT: But, again, here,

for example, if there is good evidence provided that certain epileptic medications can affect child development, et cetera, right -- so, the question is, what is the evidence we are looking for, right? If there is good evidence that there are problems if you don't fully understand how treatment of epilepsy can affect child development or can affect your health, right, that is evidence A, and there is a pool of evidence that they provide on that point.

And then, B, there is evidence that women are not understanding adequately enough of childbearing age what their options are. And the question is, what kind of evidence are we looking for to affect whatever healthcare in this regard?

CO-CHAIR TIRSCHWELL: So, you know, I think the evidence that we are looking for in this criteria, the best evidence that would be available would be if there had been a randomized trial of discussing this with

pregnant women and it led to less malformations as a result. That would be top of the line.

What is the case for all of these measures, including a number that have failed already, is that there is lots of evidence that treating sleep disorders or depression or getting rehab therapy in Parkinson's disease is beneficial, but not that that measure, as it was constructed, is going to lead to all those better outcomes. And it is that lack of linkage which has been, I think, the issue with the other measures and probably continues to be the issue to some degree with this one.

Salina? And then, Bill and Ramon.

MEMBER WADDY: It seems to me, I mean, are there other concrete measures that have been developed or is this just the very first attempt at pregnancy in women, such as use of folate or developing a requirement that they develop a strategy in case the person becomes pregnant, so that they understand it,

rather than your just having this open-ended, not open-ended, but sort of random conversation?

CO-CHAIR TIRSCHWELL: Right. So, again, you are sort of bringing up the point that has come up in related ways. Is there a way that we can get closer to valuable actions as opposed to just the discussion with the assumption of an action down the line.

Bill, were you next, I think?

MEMBER BARSAN: Yes, the only other thing I was going to add to that is I think you could actually move this a lot closer from the assessment to something meaningful if it were a very simple thing. That is that evidence that once a year in any woman of childbearing potential they are asked if they plan on becoming pregnant in the next year. If you ask that question alone, you would open the whole topic of pregnancy, and whatever. If there were questions about that, I think it would at least get the discussion

started.

As it is, it is a little bit nebulous as to what the counseling is. It is not real clear how you measure it.

CO-CHAIR TIRSCHWELL: Ramon? And then, Peter.

MEMBER R. BAUTISTA: There is about a 91-90 percent chance that pregnant women with epilepsy are going to have normal pregnancies anyway, no matter what you do, as opposed to 98 percent chance of the average person without epilepsy. So, the effect size is really very small.

In other words, even in the best of circumstances, the difference between the morbidity rates for those with epilepsy and without epilepsy is still going to be very small, and that is where the difficulty is.

CO-CHAIR TIRSCHWELL: Peter?

MEMBER SCHMIDT: So, quickly, to your comment, if you reversed that and made it an odds ratio, it would be pretty dramatic.

In the issue of RCTs, we should be accepting things like all-or-none evidence as valid. RCTs are only really done by pharmacies, pharmaceutical companies and the NIH. There are other levels of evidence that are just as compelling.

I think that we could address this. This could be assessed as an all-or-none-type criteria. If we are saying that everybody who has this, if we are saying everybody, we should be counseling everybody, and by this measure, we are pushing people to counsel everybody. We are defining counseling everybody as quality care. That counseling has evidence that it has an effect.

You don't need a randomized trial. That fits the all-or-none criteria, and so it can be considered valid. So, you don't have to go to an exception in a case like that.

CO-CHAIR TIRSCHWELL: Raj?

MEMBER SHETH: The AAN actually has practice parameters that address this

issue. The question is, does the intervention that they suggest actually affect outcome? That link is the one we are debating at present.

But there are several levels of evidence that are below that, particularly with regards to the malformation rate that is very clearly defined. We do know, for instance, that the malformation rate with valproic acid is somewhere in the order of 15 to 20 percent of all pregnancies.

We do know that low dose versus high dose affects the impact as well. We also know relatively sure, not proven, that administering folic acid does not reduce the risk of valproic-associated malformations. So, there are several pieces of evidence that are under the surface that are known, but they haven't really come up to the surface with the developers' recommendations here.

CO-CHAIR TIRSCHWELL: Any others? Salina, do you have another comment?

MEMBER WADDY: Yes, the only point I wanted to add is regarding Bill's comment, if you plan to get pregnant, and I do think that that is important because that can stimulate a conversation, but I am not sure how many people who have epilepsy or on these medications have planned pregnancies versus not, if they haven't had the conversation.

CO-CHAIR TIRSCHWELL: Terry?

MEMBER RICHMOND: Yes, I have had trouble. The last time I had trouble about education and counseling things, and we have had that discussion here. I agree that the evidence really isn't there.

However, I am more favorable to this in the sense that there is clear evidence that things can hurt the woman and fetus, you know, that piece. Just as counseling, is there evidence for counseling?

And while that is not there, I am with you on the exception thing, I think, in that this is at least a very specified

population for a very specific thing where we know harm can be done. So, I think I look at this in a very different way that sort of a generic discharge teaching, for those reasons.

CO-CHAIR TIRSCHWELL: David? And then, Ramon.

CO-CHAIR KNOWLTON: I completely agree with you and some of the other comments. I think on this one we have a lot of measures like this that are used, required counseling that we measure people who have HIV regarding safe sex. We have smoking cessation. We have some genetic disorders where we have genetic counseling in terms of childbearing years.

We have seen this from the health plan side, some problem getting people covered because they need different diagnostic testing. And there is quite a bit of ignorance that surrounds epilepsy and pregnancy. And so, I agree with exactly what you said, Therese. I think that I feel differently about this measure than I do other

ones, and there is harm done when people don't get the type of information that they need in this.

CO-CHAIR TIRSCHWELL: Ramon?

MEMBER R. BAUTISTA: Yes, my comment wasn't meant to dissuade this measure. In fact, it was actually to just point out that it is hard to get evidence for something like this. It is very hard.

At this point in time, I mean, prescribing folic acid, for example, for women of childbearing age with epilepsy is actually standard of care. It is not even a question that we ask ourselves in this day and age. It is very hard to get the evidence that we really normally look for for this kind of a question.

CO-CHAIR TIRSCHWELL: Salina? Then, Raj.

MEMBER WADDY: Well, my main issue is, I don't have a problem with having this type of measure. I think it should go a step

further where you have to have a documented plan within your chart and show that you have discussed it or given it to the patient as well, and not just check off a box that "I talked to them." You don't know so much what is involved in that conversation.

CO-CHAIR TIRSCHWELL: Raj? And then, David.

MEMBER SHETH: There are issues here that can be easily addressed. One, for instance, is what is the impact of breastfeeding on the fetus if a mother is taking medication and has epilepsy? That has a huge impact, and there is a lot of data that is out there that can be formulated into a plan of action and can affect outcome, I think.

I think the relationship is two ways. I think the relationship is, what happens to the pregnancy in a woman who becomes pregnant? And then, the second issue is, what happens to the epilepsy in a woman

that becomes pregnant? So, it is bidirectional.

You can have seizure control that is completely out of whack when you are pregnant because of blood volume changes or medications or the false belief that the moment the mother knows that she is pregnant, she stops the medication, when, in fact, the teratogenic effect has already occurred by the time the pregnancy test is confirmed, because it is in the early month of pregnancy.

So, I think this is definitely a measure that has a very defined population, as you well said, Therese, and I think it really needs rework on that.

CO-CHAIR TIRSCHWELL: Michael? Then, Salina.

MEMBER KAPLITT: Yes, I think if you look at the numerator statement, where what they are measuring here is counseling women specifically about how epilepsy and its treatment could affect contraception and

pregnancy.

If we all agree that the evidence is there, and much of it is in this document, supporting the idea that epilepsy treatment can affect contraception in pregnancy, and if there is good evidence provided here, which I think there is, that a large percentage of women feel they are not getting that information, so they don't know, then I think that, even in the absence of a randomized controlled trial, this is one of those again yes/no things, where if women don't know what this can do, and we know that this can harm pregnancies and contraception, then it absolutely has to change care.

Because if they don't know, and you are informing them, and we know that that information is relevant, then I think this is more than just evidence from an expert panel. I think that this rises to a different level than some of the evidence that we have considered earlier.

MEMBER SCARIANO: Yes, I think that this is actually really important. On the patients I have who have seizures, you know, I often look and see what medications that they are on. I once had an OB doctor tell me that, "I have your patient here, and she wants to get pregnant and she is on Depakote." I said, "Yes, but she is controlled on Depakote." He said, "Well, she is not going have an OB doctor."

So, I mean, you have to plan ahead and see if someone wants to have children. You have to think about what pills that they have to be on and actually tell them that, "As soon as you are pregnant, tell me. If you find out that you are pregnant, actually don't stop the pills" -- I just think it is actually really important.

MEMBER R. BAUTISTA: I mean, it sounds like counseling is really only a surrogate measure here. The real measure here is interventions for women who might become

pregnant. I mean, you know, not prescribing cytochrome P450, for example, if you use birth control pills or using folic acid, for example, if you intend to become pregnant. The counseling thing is really a reflection of all that, I think, that really takes place as part of patient care.

CO-CHAIR TIRSCHWELL: David?

CO-CHAIR KNOWLTON: I just didn't want to miss Raj's point because, when you read the numerator again, this says the impact, the concept about epilepsy and how its treatment may affect contraception and pregnancy.

It also, to Raj's point, it can also affect the treatment of the epilepsy. That is a risk factor, and it is something that should be discussed with a woman.

To Salina's point about a plan, a plan is important, but in this particular case the patient is a real party to that plan. That plan can be blown up pretty easily by the

patient saying, "I want to do this anyway."

And so, this is a very, very complicated issue. That is why I think it is so important. Just for the note of the developer, there is an impact on the epilepsy treatment as well with the pregnancy, and that is not in the numerator statement.

CO-CHAIR TIRSCHWELL: Jolynn?

MEMBER SUKO: This may be from just a non-clinician in the room, but I think this is bigger than just pregnancy. It also affects the choices that women would make probably about the type of contraception they would use. And so, this has a huge impact, based upon what I read in the specifications. I hear everybody talking about pregnancy, but I think it is much bigger than just pregnancy.

CO-CHAIR TIRSCHWELL: Salina?

MEMBER WADDY: Yes, so I agree with all three of those points and actually think the numerator should probably be expanded to include those.

CO-CHAIR KNOWLTON: I don't think they would have to do that to have us vote on it. I think they are hearing the debate.

MEMBER KAPLITT: I mean, the discussion here is, is there evidence to support this measure with this numerator. Right now, there may be a lot of other things in the world we could do. But what I am asking is, is there a negative to this numerator as it is written, which is the discussion in hand here?

CO-CHAIR KNOWLTON: Right, right. Yes, that was my point. I think they are listening, and they might say nobody is going to object to expanding it; we understand it was an oversight.

CO-CHAIR TIRSCHWELL: Do you guys want to ask a question to the developer or not?

All right, Raj, maybe the last comment.

MEMBER SHETH: Yes, I think what

might be important for the developer to do is actually broaden the degree of support, bring in other organizations that might help with the development. I know that the Epilepsy Foundation is probably another critical element in this that would be very interested, has a vested interest in serving this community as well. So, it might be broadening it would be an option, too.

CO-CHAIR TIRSCHWELL: Yes, do ahead.

MS. SWAIN-ENG: That is just what I was going to say. We did have the Epilepsy Foundation of America that was involved and the American Academy of Family Physicians, the American Academy of Pediatrics, numerous different health insurers, NAAC. All those groups were involved.

CO-CHAIR TIRSCHWELL: We just didn't know that because it wasn't written down, I guess, right, Raj?

MS. SWAIN-ENG: No, it is in

the --

CO-CHAIR TIRSCHWELL: It is? Oh, it is not in the summary we got. Okay. Thank you.

So, let's go ahead and vote on the evidence. You could either think there is sufficient evidence -- if you want to go with the exemption, which has been brought up on a couple of occasions now, then you need to vote for No. 2, is that right? And if neither of those, then I guess three.

Yes, let's go.

(Vote taken.)

MS. THEBERGE: I need two more responses. One more. Is anyone missing? Can everyone vote one more time? There we go.

Eleven yes; 13, no, evidence does not meet guidance.

CO-CHAIR TIRSCHWELL: So, I think that means that we need to vote, then, on the exception rule.

Does anybody have any other

comments they want to make about the exception to empirical evidence before we vote?

(No response.)

Okay. Let's go ahead and vote on this then.

(Vote taken.)

MS. THEBERGE: I need one more response. There we go.

Twenty-three yes, 1 no.

CO-CHAIR TIRSCHWELL: All right. We are past that hurdle.

Raj, now we want you to briefly discuss high impact.

MEMBER SHETH: I think that some of the impact has already been discussed.

CO-CHAIR TIRSCHWELL: Okay.

MEMBER SHETH: But the impact, obviously, from a population number, this is a big impact. We are talking about half the population with epilepsy could potentially be affected, obviously excluding the younger children and those over 44 for the women.

So, it is a big impact factor. The consequences of not getting the advice, not understanding their risks, really has an impact on the fetus. That is one. And it may have a lifelong impact on the patient. So, clearly, a very high impact.

CO-CHAIR TIRSCHWELL: Great. I think, yes, let's go ahead and vote.

(Vote taken.)

So, no issues there. And then, 1(b), the performance gap or the opportunity for improvement.

MS. THEBERGE: I just need to read out the numbers for the transcript.

CO-CHAIR TIRSCHWELL: Sorry.

MS. THEBERGE: Twenty-three high, 1 moderate.

CO-CHAIR TIRSCHWELL: Just trying to keep the train rolling.

(Laughter.)

Raj, performance gap.

MEMBER SHETH: The performance gap

I think has been established. Clearly, studies vary between 2 and 20 percent of women are counseled with regards to their epilepsy risk. So, this doesn't even hit the 50-percent mark. So, clearly, there is a performance gap.

CO-CHAIR TIRSCHWELL: I think that is good enough.

(Laughter.)

Let's go ahead and vote.

(Vote taken.)

MS. THEBERGE: Twenty-four high.

CO-CHAIR TIRSCHWELL: Make a note. All right. Now we are on to scientific acceptability, reliability and then validity. But they are combined here because this has not been out before, right? It has not been tested.

Okay. So, reading out the slide, the reliability part is for the specifications. They are unambiguous, likely to consistently identify who the population

is, identify the process, and compute the score, and that the specifications also reflect the quality-of-care problem and the evidence that we have ignored.

MEMBER SHETH: So, I think here there is very little controversy. I think on both counts the population is clearly identified. I think that the evidence that exists is quite high.

CO-CHAIR TIRSCHWELL: And it will identify the problem at hand?

MEMBER J. BAUTISTA: I would disagree.

CO-CHAIR TIRSCHWELL: Okay.

MEMBER J. BAUTISTA: Yes, I think the measure specifications are not at all precise. I mean, the scope is huge. It is impact of epilepsy on contraception and pregnancy. What is exactly meant by "counseled"? There is no operational definition.

CO-CHAIR TIRSCHWELL: Okay.

Salina?

MEMBER WADDY: I agree. I mean, that is the issue that I have been having with this measure. If it can be more specific or somehow, even if the AAN had some type of structured basic conversation to have that was required, that would make it much simpler. But this is very open-ended.

CO-CHAIR TIRSCHWELL: Terry?

MEMBER RICHMOND: I am usually really into preciseness here, but I am not sure how much we could micromanage this. Because how we would talk to a 12-year-old who just sort of could potentially be pregnant versus a 30-year-old versus a 40-year-old, and the issues we would counsel about I think would be really different.

So, in terms of really specifying at a high level, it just does not ring true to me as a clinician. So, I understand the concerns, but I am not sure how we would deal with that.

CO-CHAIR TIRSCHWELL: Jocelyn, do you still have more?

MEMBER J. BAUTISTA: Well, I think maybe, then, that speaks to whether this really meets NQF criteria. I mean, this is good standard of care. I don't argue that this is very important to do in your day-to-day work, but does it meet criteria if we are not able to have precise measure specifications?

CO-CHAIR TIRSCHWELL: Michael, Salina, and then Peter.

MEMBER KAPLITT: Yes, I mean, I am just wondering how you would capture that because there is a CPT code, right, for this, which is counseling women of childbearing, or whatever, but it is basically this thing.

So, I am wondering, separate from anything else, if you make this too specific, how exactly are you going to make this useful and capture it? That is my concern.

CO-CHAIR TIRSCHWELL: Salina?

MEMBER WADDY: I agree with you, but I am not talking about developing some type of script, which I think would be completely inappropriate. But, basically, discussing the medications that have the highest fetal anomalies as well as the impact that it could have on your type of epilepsy, potential decisions regarding avoidance of pregnancy and what the options are, and the use of folate. I mean, those just very basic things, but within this, if you were talking about something like driving, it wouldn't be enough to say, you know -- I don't know how to answer that beyond that.

CO-CHAIR TIRSCHWELL: Peter?

MEMBER SCHMIDT: So, before coming here, I reviewed a number of clinical practice guidelines, you know, quality indicators that have been very successful elsewhere. And they are not that specific. You have to allow the clinician to make choices about how to address something.

And so, with the depression in Parkinson's disease, there was evidence that depression is difficult to diagnose in Parkinson's disease. So, that is something where we have evidence to back up a request for specificity. But if we don't have evidence that addressing this issue is challenging, then it is difficult for us. We should not apply our own opinions about that some other clinician is going to fail at doing it, just because we don't trust them. You know, you have to let the clinician have some autonomy.

CO-CHAIR TIRSCHWELL: Any other comments about the reliability or validity, really mostly related to the specifications of this measure, before we go ahead and vote?

(No response.)

Okay, then, I think we should go ahead and vote.

(Vote taken.)

MS. THEBERGE: I need three more

responses.

CO-CHAIR TIRSCHWELL: There you go.

MS. THEBERGE: Twenty-one yes, 3 no.

CO-CHAIR TIRSCHWELL: All right. So, we are in the relatively-unchartered territory, usability.

(Laughter.)

Raj, comments on usability?

Risha, could you take over his microphone, please?

(Laughter.)

MEMBER SHETH: From the general feeling of usability, the group felt that there was a high degree of usability for this.

MEMBER J. BAUTISTA: Again, I disagree. There is no data at all about usability. I mean, how can you judge? There is no data submitted.

MEMBER SHETH: The data is not submitted, but it clearly exists.

CO-CHAIR TIRSCHWELL: So, there is a CPT code. No data is reported, but, apparently, the AAN is using it.

Why don't you guys comment for a moment about the usability?

MS. SWAIN-ENG: So, the measure is already in use by multiple different programs. It is in use in our neuro-protective program, which, again, our maintenance and certification Part V program. The feedback that we have gotten from the clinicians, we have had 119 who have purchased the epilepsy module, which includes this measure, and have had no issues specifically with this measure.

This is one of the measures they find to be the most helpful, that has really helped them improve their practice, really brought a sense of awareness to them, things that they hadn't considered, that they needed to counsel a patient who was young about possible contraception issues, you know, different things that they hadn't considered

before. So, this has really helped them and they haven't had issues.

As Dr. Bever was mentioning, too, this measure is also in the PQRS 2012 program, and we have a registry through CECity that is approved. It is a CMS-approved registry where patients can report on this measure that will go directly to CECity. So, we are starting to aggregate a little bit of data from that. It did just open in August, and we have 11 people who have enrolled in this program so far.

So, we are definitely accumulating data and haven't heard any issues with usability at all.

CO-CHAIR TIRSCHWELL: And is this one up for time-limited?

MS. SWAIN-ENG: Yes.

CO-CHAIR TIRSCHWELL: So, then, we will hear back, or at least the NQF will hear back with some data in a year's time. And hopefully, you will be able to close the loop a little bit on some of these issues.

Any other comments on usability?

Salina?

MEMBER WADDY: Yes. So, I just want to be clear. So, physicians have told you that it has changed their practice. Has it actually improved quality of care?

DR. BARKLEY: May I make a comment, please?

CO-CHAIR TIRSCHWELL: Is somebody on the line?

Hold on one second, please.

Were you addressing that question, Salina, to the developers?

MS. SWAIN-ENG: He is part of us.

CO-CHAIR TIRSCHWELL: Okay, go ahead on the phone. Can you identify yourself, please?

DR. BARKLEY: My name is Gregory L. Barkley. I am a neurologist at Henry Ford Hospital in Detroit, and I am an epileptologist. I was involved with the committee that helped develop these.

We have an abstract that we are going to present at the American Epilepsy Society meeting this December where we actually looked at clinical documentation of patients with epilepsy seen by neurologists at Henry Ford Hospital as well as the epilepsy specialists.

For this particular question of the childbearing potential, the documentation went up dramatically in terms of our awareness of the need to do that, and the documentation of this discussion was being held with women. As others noted, this opens up a whole can of worms about will my child have epilepsy or what is the right drug, all those kinds of questions.

And so, we went from about 11 percent of the charts documenting this in women of childbearing age to 56 percent amongst the epileptologists, just in documentation. So, I am sure this has made an impact on the quality of care of these

patients.

CO-CHAIR TIRSCHWELL: Followup from Dr. Waddy?

MEMBER WADDY: Yes. So, I mean, do you have other measures like either compliance or adoption of some of the AAN practice parameters? Did it increase the number of patients who were on folate? Did it change people who were on valproic or I think Topamax? Do you have any evidence that there was actually change in the quality of care?

DR. BARKLEY: Actually, our abstract or our research didn't address that, but I am sure that, when you have these discussions, particularly we were involved with Kimford Meador's neonatal outcomes of anti-epileptic drug program, which showed, in particular, that valproic acid was negatively correlated not only with the presence of birth defects, but the intellectual outcome at four-and-a-half years, which is clearly lower than all of the other three anticonvulsant drugs in

that prospective study and independent of maternal IQ.

So, this is really, as opposed to what Raj -- I am not sure of your last name -- said. This is a third-trimester effect, more likely than the first-trimester effect for the congenital malformation. So, I think this really has the potential to change care. Once you start to discuss this, women make big changes in their decision about what they are going to do about getting pregnant, which drugs to take, whether they should be on an IUD versus an oral contraceptive medication.

CO-CHAIR TIRSCHWELL: Okay. Thank you very much.

I just will suggest that we have strayed back into evidence when we want to be talking about usability.

Peter?

MEMBER SCHMIDT: Yes, just a similar comment. We seem to be conflating usability and feasibility. Usability is

defined as for public reporting and accountability. And I think that because of the concept of the time-limited endorsement, that usability is something that we assess once there is data.

MEMBER WADDY: Yes, but the problem is 3(b). That is the one that I have. That is what was generating my questions, is whether it is meaningful, understandable, and useful for quality improvement.

MS. BOSSLEY: Right. This is Heidi.

This is one that everyone struggles with, especially when measures are not yet tested, because there is a little blurring of evidence and validity, I think.

But any new measure that comes in, there, first of all, is not an expectation that it be in use when it comes into NQF. This measure actually is in use. So, they are ahead of the game in that way.

So, what we really are asking you

to look at is, based on the information you have and what you have heard from those on the phone and here at the table, do you believe it will inform through public reporting and accountability purposes and could it for quality improvement?

We don't expect them to come back with that data until maintenance the next time. So, again, it is, do you believe, based on what you know now, that it could be useful and usable?

So, it is going to be a little vague because you don't have data yet, but that is kind of where we are now with new measures.

CO-CHAIR TIRSCHWELL: Salina?

MEMBER WADDY: I mean, at least for 3(a), to me, it seems very much absolutely; that is kind of a no-brainer for me.

But, for 3(b), because of the way it is structured -- and I don't know really

what is going to go on in that conversation, if that conversation actually changes practice, and it leads to changes. That is the one that I am having trouble with.

MS. BOSSLEY: And I think you should rate this criteria against that. I think that should be one of the factors, and one of the questions could be, at the time of maintenance, assuming this is endorsed, can AAN come back with some information on that?

CO-CHAIR TIRSCHWELL: I just want to remind people, as I was just reminded, that this is not a "must-pass" criteria.

MS. BOSSLEY: Right.

CO-CHAIR TIRSCHWELL: So, even if you vote it down, the measure can still pass, and your objection would be noted.

(Laughter.)

We're done; let's vote.

(Laughter.)

(Vote taken.)

MS. THEBERGE: I need three more

responses.

Ten high, 12 moderate, 1 low, 1 insufficient.

CO-CHAIR TIRSCHWELL: Okay. Very good.

And then, finally, we are on to category 4, which is feasibility.

Raj, do you want to address that?

MEMBER SHETH: So, I think the issue is, how would you ascertain that this has been done? Again, this would be a checkoff box, and it would be done perhaps on a yearly basis.

The implementation is unclear. You know, what do you do with paper records? Are you able to abstract that aspect of it? And the general feeling of the group was that there was insufficient data that was provided to support this.

So, the overall feeling was that it was feasible. There is a CPT code that you can look at that that would assess whether

counseling of women was done, but there were some members of the group that felt that they were not quite sure how you would collect it in paper medical record terms.

CO-CHAIR TIRSCHWELL: So, there are some details, and in a pure EHR environment with CPT codes it might be easier to describe exactly how it would all happen, but there is a little bit of fuzziness. It is in use now. We would, hopefully, have more information in a year or so.

Any other comments?

This is also not a "must-pass" criteria. So, again, even if you vote against it, it won't necessarily affect the final outcome.

Other comments?

(No response.)

Okay. Let's go ahead and vote then.

(Vote taken.)

MS. THEBERGE: We need three more.

We need one more vote.

Four high, 15 moderate, 2 low, 3 insufficient.

CO-CHAIR TIRSCHWELL: Okay. Very good. So, I think we are now on to the overall evaluation at this point.

Any further discussion before we vote on this overall?

(No response.)

Okay. Let's go ahead and do it.

(Vote taken.)

MS. THEBERGE: We still need one more response. There we go.

Twenty-four yes.

CO-CHAIR TIRSCHWELL: All right, then, moving right along to the next measure, Jocelyn, 1953, seizure type and current seizure frequencies.

MEMBER J. BAUTISTA: So, this is also a new submission from the American Academy of Neurology.

So, this measure captures the

proportion of epilepsy patients who are being seen for epilepsy for whom seizure type and current seizure frequency are documented in the medical record.

It excludes those patients who have a documented medical or patient reason for not recording seizure type or seizure frequency, such as the patient is unable or unwilling to communicate or provide that information.

And the level of analysis is at the clinician level.

CO-CHAIR TIRSCHWELL: And evidence?

MEMBER J. BAUTISTA: Evidence. So, the question is whether there is evidence that documentation of seizure type and seizure frequency leads to better outcomes. There is not such good evidence for that in terms of the documentation. But the implication is that seizure frequency is really the main outcome measure in epilepsy, right? And so,

if you don't even document it, you can't impact it.

So, the implication is you document, you ask and you document the seizure frequency, and then you are able to act on it. So, it is, again, there are multiple steps to the improved outcome.

So, we again run into this evidence issue.

CO-CHAIR TIRSCHWELL: Daniel? Then, Risha.

MEMBER LABOVITZ: I am a stroke doctor, but I have a deep love for dealing with epilepsy problems. I have looked at epilepsy classification. I cut my teeth on it in training.

It is a total quagmire.

(Laughter.)

Epileptologists are now duking it out. There is a new classification scheme that has been proposed. You may hear the roaring. Those are the dinosaurs over here

and people in spaceships over there. There is a huge fight going on about classification.

And the question is, does that affect outcome? I don't see that we can even classify epilepsy right now or at least make providers do it.

CO-CHAIR TIRSCHWELL: It doesn't say you have to get it right.

(Laughter.)

MEMBER LABOVITZ: Yes, you don't have to get it right, true, but, then, I think that begs the question of does it help.

CO-CHAIR TIRSCHWELL: Risha?

MEMBER GIDWANI: Yes, I had a similar concern. The NICE guideline says that "The established classification system is undergoing review. Current proposals have the status of work-in-progress," and that failure to correctly classify an epilepsy syndrome can lead to inappropriate treatment and persistence of seizures.

So, I think if the field as a

whole hasn't come to a consensus about how to categorize epilepsy properly, I wonder if some of the harms of this are just that physicians will now feel pressured to start classifying, use an incorrect classification scheme and then go down an inappropriate treatment pathway.

CO-CHAIR TIRSCHWELL: And any other comments?

Jack?

MEMBER SCARIANO: Yes, well, if you have an actual focal epilepsy, that always makes me look harder, and it also may make me look and get more MRI scans over a period of time. So, if you have focal epilepsy or if it is just unilateral onset, there is a possibility that even epilepsy surgery may help. So, if you have a focal epilepsy, I think it is really important to actually document that.

CO-CHAIR TIRSCHWELL: Yes, I mean, I would just add that, despite the fact that

the classification systems are under discussion, describing the types of seizures the patient is having, even just in plain English terms, and the frequency with which they are happening, seems like a pretty minimal standard of care for an evaluation, especially in a neurology clinic, for anybody that is being seen with epilepsy.

Ramon, and then Risha, and then we will get to you guys over there.

MEMBER R. BAUTISTA: Yes, actually, we are talking about seizure types right now, not epilepsy classification. That is our next discussion, actually.

But, going back to your comments, I agree, David, that for the most part we know how to at least think through epilepsy and think through seizures, enough for us to make any significant change in the way we manage them. So, I don't think it is a big issue.

CO-CHAIR TIRSCHWELL: Thank you.

Did somebody else have their thing

up? Go ahead, AAN.

DR. BEVER: So, the working group that came up with this measure was motivated by the fact that the drugs are tested in specific subtypes. They acknowledge the fact that in details there is a lot of controversy about the classification of different seizure types, but, broadly, there are some large groups that do relate to the appropriate anticonvulsive medication that should be used in the patient.

And there was felt to be a gap at least in some providers in terms of their understanding of the patient seizure type, and based on referrals to epileptologists, a lack of documentation of a seizure type that would lead to a proper selection of a medication. So, there was felt to be a gap in care, and that you could not choose proper medications without actually identifying the seizure type, at least in terms of the drugs that you were choosing among. So, that is how they came up

with this.

CO-CHAIR TIRSCHWELL: I mean, as you are describing it there, it begs the question for the next measure about overlap. We can get to that when we get to the next measure.

Risha, did you have something different to add?

MEMBER GIDWANI: No, just the same point. I think we are conflating epilepsy with a seizure. So, if we could just stay on the epilepsy component right now?

CO-CHAIR TIRSCHWELL: Seizure.

MEMBER GIDWANI: Aren't we doing epilepsy at the moment? Then, my fault. I am sorry.

CO-CHAIR TIRSCHWELL: It is the diagnosis of epilepsy, but it is the seizure types that they are having.

MEMBER R. BAUTISTA: I mean, just for education for the group, just to make sure you understand the difference --

CO-CHAIR TIRSCHWELL: Please.

MEMBER R. BAUTISTA: -- when you classify or diagnose seizure types, you refer to things like the localization of the seizure. Is it a generalized or a partial seizure? Is it a temporal lobe or frontal lobe seizure. And you also refer to the clinical semiology? Are you dealing with a generalized tonic-clonic seizure or are you dealing with a complex partial seizure, or an abson seizure?

Epilepsy classification, on the other hand, refers to the classification of different diseases that cause seizures. So, for example, you have something called idiopathic epilepsy, cryptogenic epilepsy, symptomatic epilepsy. That is how you distinguish between seizures and epilepsy. One is a disease-specific diagnosis; one is a characterization of what goes on during the seizure.

CO-CHAIR TIRSCHWELL: So, any

other questions?

Yes, Risha, go ahead.

MEMBER GIDWANI: Just for the record, I will withdraw my previous statement and apply it to the next measure then, when we review that.

(Laughter.)

CO-CHAIR TIRSCHWELL: Thank you for making that official.

Okay. So, let's go ahead and vote on the evidence for this measure.

(Vote taken.)

MS. THEBERGE: I need one more response.

Yes, 11; no, evidence does not meet guidance, 9, and then 4, no, insufficient information.

CO-CHAIR TIRSCHWELL: Okay. So, let's go back to high impact.

(Chorus of noes.)

Oh, I'm sorry, I was just looking at the size of the bars there.

(Laughter.)

Okay. There you go. So, then, moving along to the other Dr. Bautista, 1954.

Are you guys related, by the way?

MEMBER R. BAUTISTA: All right. So, let's talk about 1954. So, 1954 actually documents etiology of epilepsy or epilepsy syndrome. So, the denominator is these are the patients with a diagnosis of epilepsy, and the numerator states at least documenting the actual epilepsy classification or syndrome.

In other words, you want to write down if they have cryptogenic epilepsy or symptomatic, and you might want to be more specific. Do they have post-traumatic epilepsy, and so forth and so on? Or do they have idiopathic epilepsy? So, try to make the orderly diagnosis of patients you see every time you see them.

Let me put on my schizophrenic hat here because I do have mixed feelings about the measure which I will try to explain.

No. 1, the measure is supposed to be used not just by specialists, right, but also by the general doctors. Okay, good. And that is one problem I have with this measure, is that I am not sure a non-specialist or a non-neurologist would be in a position to actually make the proper classification of epilepsy syndrome or epilepsy type.

Secondly, as far as the evidence is concerned, they actually point out both the SIGN and the NICE study, both of which are really, if you look at it, position papers. They don't really give details on this, on the necessity to put down the epilepsy syndrome.

On the other hand, there are tons of evidence out there that link particular syndromes to different treatment options. For example, we know that mesial temporal sclerosis is linked with epilepsy surgery. We know that idiopathic generalized epilepsies have a certain select number of drugs that you can choose from. So, this is all out there.

It is not just documented in the literature as it is.

Furthermore, the actual SIGN and NICE study actually documents early on that epilepsy has to be diagnosed by a neurologist or an epileptologist. So, in a way, choosing you want to hear from the SIGN and NICE studies, but choosing to dissuade what they don't want to hear, and that is a problem I have.

So, my main point is that although the papers as written do not provide good enough evidence, from the literature there is tons of evidence that actually suggests the importance of proper documentation of epilepsy syndrome.

CO-CHAIR TIRSCHWELL: Anybody have any comments on this particular measure for the evidence base?

(No response.)

So, let's go ahead and vote on it then.

Oh, I'm sorry. Daniel?

MEMBER LABOVITZ: I was just going to say I think we heard from the lesser Bautista about the lesser measure.

(Laughter.)

This one is even more fraught than the one we heard before. Epilepsy classification, really, I would say right now hopeless. Seizure classification, bad; epilepsy classification, hopeless.

And it just makes it very hard. There is clearly a role, and epilepsy doctors work very hard to choose drugs appropriate to the disease. And there are some epilepsies which require specific drugs. That is the role of the specialist.

But I think asking the primary care doctor to get this right, and then to make the right choice, when the specialists can't agree, is a hopeless prospect.

CO-CHAIR TIRSCHWELL: I guess I have a question, and anybody can answer this.

I don't know the answer myself.

When a primary care doctor sees a patient for one of these neurological syndromes, do they write on their billing codes only the things that they are really steering the ship for, the hypertension and the diabetes? Or do sort of all of the patients' diagnoses get bundled in because more diseases, higher coding, better reimbursement. Who knows what the motivation for that is? Does anybody know the answer to that?

MEMBER WADDY: No, that is why I brought that up about Parkinson's, that they may see them for their problems with eating or something, but somehow bundle that in. How accurate really does that reflect what happens in the visit?

CO-CHAIR TIRSCHWELL: Yes. Yes, go ahead, Jordan.

MEMBER EISENSTOCK: I was just going to say I don't have any data behind

this. This is just an opinion.

But I think with the EMR and the implications of its being easier to just sort of check off all those diagnoses and they are being kept track of visit to visit and among different specialists and PCPs, that probably we would see that.

CO-CHAIR TIRSCHWELL: We would see more of it even with the EHR.

MEMBER EISENSTOCK: Exactly.

CO-CHAIR TIRSCHWELL: Jack, do you have a comment?

MEMBER SCARIANO: Yes. On the patients who I see off the primary care doctor, almost all of them who have any type of a sinigual spell may have been diagnosed as having seizures. So, yes, if they even think there is a seizure, they put it down.

CO-CHAIR TIRSCHWELL: Okay. First, Terry, then Salina and Ramon.

MEMBER RICHMOND: Yes, so the thing I got confused about this is, when

patient comes in, if they are coded for epilepsy -- so, if you have a primary care who is taking care of a stable epileptic who is managing their anticonvulsants, they probably will have a code generated. And yet, it seems to me like -- I am married to an epileptic, so I will speak as a consumer here -- so, it seems to me we know the source. He has scar tissue on his brain. His primary care manages his anticonvulsants. I am sure she probably checks the CPT code. But I don't think every time she sees him she needs to say he has a scar on his brain tissue and document that on the medical record. Maybe I am missing something, but --

CO-CHAIR TIRSCHWELL: Well, honestly, I think it should say post-traumatic epilepsy, that simple, and you have done it at that point, if that is what --

MEMBER RICHMOND: But every time, every six months, if you are seeing somebody every six months?

CO-CHAIR TIRSCHWELL: Well, yes, just that phrase is all you need.

MEMBER RICHMOND: I mean, I am just not clear on those.

CO-CHAIR TIRSCHWELL: It should be probably automatically applied, I would think.

But, anyway, Ramon?

MEMBER R. BAUTISTA: Just to answer the question about the coding, there is, I think it is an ICD-9 code for the epilepsies from 345.1 to 345.9. In the course of actually mainly a hodgepodge of epilepsies and seizures, there is a catchall code, though, 780.39, which actually is an epileptic-seizure-type code. So, in other words, to answer your question, the primary care doctor has a way of having a catchall code for all of these.

CO-CHAIR TIRSCHWELL: And many of these EHRs list your problems by an ICD-9 code, and it is actually included in your next whatever.

Salina? And then, Michael.

MEMBER WADDY: That was one of the things that I was thinking of as well. Certainly, in a physician's office, what you don't want is for a person to go like 10 years and it hasn't been updated. And so, I think it is a little bit better if you carry those forward.

But my actual question is, what is this really trying to accomplish? I mean, at the end of the day, are you just trying to document how well they do this or are you trying to match are they prescribing the medication that is appropriate for that syndrome, and if so, then that really should be the measure instead of this.

CO-CHAIR TIRSCHWELL: So, again, Dr. Waddy brings up the point, is this too far back in the chain of events to necessarily cause the improvement in outcomes and quality that we are looking for?

Michael?

MEMBER KAPLITT: To that point, I mean, putting aside the poor primary care physician that has gotten horribly brutalized here today -- (Laughter) -- you know, the numerator, as was said earlier, is every single visit that this is documented and reviewed, right, at each visit? So, the question is, what evidence is there that that does anything? Is there evidence that this is something that is changing, that requires this to be reviewed, that the diagnosis is changing, requires it to be reviewed every time? Is there evidence that that does anything?

And the reason that matters, on top of everything else, is that we have all been hearing lately now the government is starting to go after cloned notes, right? Well, we are promoting cloned notes here by saying you are going to do the same thing every time, even though it is not changing. We are just going to be encouraging people to

just cut and paste the exact same thing every single time, every note for 10 years. So, what is the evidence that it is going to change anything?

CO-CHAIR TIRSCHWELL: And, in fact, I mean, compared to the Parkinson's disease, which progresses and changes over time, it seems like there would be even less cause here if they have seen a specialist and gotten a good diagnosis.

Dr. Waddy? And then, Jolynn.

MEMBER WADDY: Can we just ask the developers what you wanted to accomplish with this?

MS. SWAIN-ENG: So, they are reviewing and documenting etiology of epilepsy or epilepsy syndrome with the patient at every visit. You should have gotten this document. So, I apologize if you didn't.

The clinician can determine the appropriate treatment, understand the expected response to treatment, and provide appropriate

content for counseling the patient. The outcome for the patient is better symptom management, appropriate treatment, and improved quality of life.

This measure may also lead to a reduction in overuse and misuse of treatments because etiology of epilepsy will be reviewed and documented at every visit.

MEMBER WADDY: Right. I mean, I understand that.

I jumped ahead.

CO-CHAIR TIRSCHWELL: No, that is okay.

MEMBER WADDY: I understand that; I just don't understand why the measure is not measuring -- it doesn't seem like the measure is actually measuring that part of it, the quality of care that is delivered.

MS. SWAIN-ENG: What exactly would you have us measure?

(Laughter.)

MEMBER WADDY: Well, if they have

generalized epilepsy, are they taking an appropriate medication for generalized epilepsy?

MS. SWAIN-ENG: So, you would have us develop separate measures for every possible etiology, just so I am following you?

MEMBER WADDY: I am not saying how you should develop it. It is just I think it gets back to the issue of, is it closely linked to quality of care? And this isn't measuring that, I don't feel like.

CO-CHAIR TIRSCHWELL: Okay. Jolynn?

DR. BARKLEY: May I make a comment?

CO-CHAIR TIRSCHWELL: All right, go ahead.

DR. BARKLEY: This is Gregory Barkley again.

One of the thinkings behind this is that people have talked about having specific syndromes where you expect good

outcome, for example. When they come back and you ask questions about their seizure frequency and their side effects, their medication, and they are not responding, then it challenges whether you have the correct diagnosis or the right syndrome. And then, that may lead to different kinds of diagnostic testing, and then other interventions to try to improve their outcome.

CO-CHAIR TIRSCHWELL: That is the seizure type and frequency measure, it would seem, and now we are talking about the epilepsy etiology and syndrome, which, again, it appears that there is overlap.

So, let's go to the group, a couple more comments.

Jolynn?

MEMBER SUKO: This is just more a practical comment. I think from a claim's perspective, on the physician side there is not that many diagnosis codes. So, if I was going to my primary care physician, I would be

having to go for treatment of epilepsy, and that would have to be coded on the visit.

And again, I don't think this is going to change the outcome, but just from a practical perspective, there were some questions about the coding. I think that it would be seen in a single -- I would have to be going to see you for my epilepsy, not my diabetes, and it would be that visit of epilepsy that would be counted in this.

CO-CHAIR TIRSCHWELL: I apologize, I don't know this. Is it just the primary diagnosis code that is being used for this measure or any of the diagnoses that are recorded? It is primary? Okay. So, that probably would mostly limit it to specialty care.

Michael?

MEMBER KAPLITT: Okay. So, again, I would like anybody in this room or on the phone to answer, because we are in the evidence section, to answer the following

question for me: what is the evidence that reviewing -- it is nice, the idea and the concept -- what is the evidence that reviewing the epilepsy diagnosis at every single visit changes anything? Before we get into any other discussion, I would like anybody in the room or anybody on the phone to answer this before we drift into anything else.

DR. BARKLEY: This is Greg Barkley again.

What I would say is that, if you blithely assume that you have made the right diagnosis and that you have thought that this person has a focal epilepsy, and they really have a generalized epilepsy, or vice versa, that if you don't question -- if someone comes in and is doing well, there probably isn't any evidence to need to make much of a change. But if they are not doing well, then that raises the issue, do you have the right diagnosis?

MEMBER KAPLITT: With all due

respect, that is an opinion. What is the evidence?

DR. BARKLEY: Well, there is evidence of the diagnosis of juvenile myoclonic epilepsy, which is a syndrome that comprises about 8 percent of the people with epilepsy. It is easily diagnosed if you know the syndrome. And if you don't, you end up putting the people on the wrong medication. So, knowing the syndrome and putting them on the right medication improves outcome.

CO-CHAIR TIRSCHWELL: Ramon?

MEMBER R. BAUTISTA: I would submit that, if you are smart enough to know how to classify epilepsies, you are probably smart enough to know what the treatment options are. I am not sure having to document that every time is the way to go. I think it might be more important to, I guess, show that at least you are treating them the right way. I mean, if you know how to classify epilepsies, you know what to do. That is part

of why you classify epilepsies in the first place.

CO-CHAIR TIRSCHWELL: And I apologize for prematurely going back to the comparison with the other measure, but it seems like I am hearing from multiple people that it is most important for all of this in the patients who are not responding to therapy. And so, maybe the measure with the seizure descriptions and the frequencies would be more likely to impact quality of care than the description of the syndrome.

Raj, do you have a comment?

MEMBER SHETH: Well, I think that the measures as they stand obviously suffer from all the criticisms that have been offered here. But I think there is another aspect to it that perhaps hasn't been addressed, and that is that, if you diagnosis a patient with having temporal lobe epilepsy, for instance, and you know that the evidence suggests that they are not likely to respond to medication,

you would sort of move to the next step, which would be a surgical option.

I think there is a lot of benefit. What typically happens in practices is they document seizure disorder and give them a visit to see them in six months' time, instead of actually looking at other options that might be available.

So, I think it is very important -- this is one of the AAN quality measures -- that you inquire of the patient as a surgical candidate, precisely because of this, because we know that the likelihood of remission with medication, with more medications, is on the order of 2 percent. The likelihood of being seizure-free with surgery is somewhere on the order of 70-80 percent.

So, I think if the measure were modified some, it would have value.

CO-CHAIR TIRSCHWELL: Any other comments?

Salina?

MEMBER WADDY: Yes, and I agree with you that, if it is either to assess whether or not they have the correct syndrome and they are on the correct medication, then having one measure for that. Or if it really is, as the person on the phone is saying, for you to really think about those patients that have uncontrolled epilepsy, then I think it would be more valuable to put within the numerator patients who have greater than a seizure frequency of three or some basic number over "X" period of time, and then what needs to be done.

DR. BARKLEY: May I make a comment?

CO-CHAIR TIRSCHWELL: Yes, go ahead.

DR. BARKLEY: I agree with that. Actually, for the patient, it is a very simple proposition. If you are seizure-free, you have good quality of life. If you are having any seizures, you have poor quality of life.

And so, the patient-centered measure is zero for seizure count since your last visit.

CO-CHAIR TIRSCHWELL: Okay.

DR. BARKLEY: There is plenty of evidence that shows and lots of quality-of-life studies that show that that is really the only thing that counts to the patient.

CO-CHAIR TIRSCHWELL: Okay. Yes, one more comment from the AAN.

MS. SWAIN-ENG: Just quickly, just to respond to Dr. Raj's comment about referral for surgery, we do have a separate measure that we will be bringing back to NQF. It is not currently in the PQRS program. But it is focused on patients with a diagnosis of intractable epilepsy and referring them for evaluation for appropriateness for surgical therapy.

There is evidence that shows, on average, people have a 20-year wait before they are actually referred for surgery evaluation. So, just to answer the question,

that doesn't relate directly to what we are talking about now, but just to let you know that we will be coming back to NQF. If you are on the Steering Committee again, you may be seeing that sometime soon.

CO-CHAIR TIRSCHWELL: Thank you.

Okay. Any other comments?

(No response.)

Let's go ahead and vote then on the evidence for this measure.

(Vote taken.)

MS. THEBERGE: I have 20, 21, 22, 23. I need one more response. Could everyone vote one more time, please? Nobody has stepped out of the room, right?

CO-CHAIR TIRSCHWELL: Oh, there it goes.

MS. THEBERGE: Okay. There we go.

(Laughter.)

All right. Zero yes; 15, no, evidence does not meet guidance, and 9, no, insufficient.

CO-CHAIR TIRSCHWELL: So, as we did not hear the exception brought up, I think we are done with this measure, too, then.

And on that note, should we open it up for public comment? So, should we talk to the operator?

Arnika, could you please open the phones for any public comment?

THE OPERATOR: Yes, sir.

At this time, if you would like to ask a question, please press *, then the number 1 on your telephone keypad.

(No response.)

And there are no questions at this time.

CO-CHAIR TIRSCHWELL: Any other comments here?

(No response.)

MS. JOHNSON: Okay. Thanks, guys. We have had a very interesting day one of our Phase II. So, thanks for all the thought and effort that you guys have put into this.

I am going to ask Suzanne here in just a minute to make sure I haven't forgotten anything.

But I think the one thing that I do want to remind you of is we will be spending some time tomorrow afternoon discussing the CMS Yale readmission measure. Again, the mortality measure was withdrawn, but the readmission measure is still on the table.

And to that end, I have a little bit of homework for you. I want to ask you to take a look at the comments and the responses that came in on those measures. We had already put those up on SharePoint. To make things a little easier for you, we basically put the same thing up on SharePoint, but with only the stuff relevant to the readmission measure. So, that way, you don't have to plow through. Just look at the stuff on the readmission measure and just make sure that you have had a chance to see the developer

responses.

And then, tomorrow afternoon the developers will be here and I believe are going to show you a few slides as well. We are going to allow them to do that.

So, if nobody has any questions or concerns, including Suzanne --

MS. THEBERGE: Two quick things. I just wanted to let you all know I emailed you an updated Excel sheet and Word document this afternoon that has just the comments and responses for 2027.

And then, on a housekeeping note, I have just been told our building is on lockdown because the Occupy protest is like a block away, and I guess they are right around here. So, if you need to leave -- (Laughter) -- just be aware of that, but you won't be able to get back in if you leave because you don't have a key. So, don't try to come back.

(Laughter.)

CO-CHAIR KNOWLTON: Including

through tomorrow?

(Laughter.)

MS. THEBERGE: I believe you will be able to get in tomorrow morning.

So, if you forget something, you will just have to probably get it tomorrow morning, since you will need a key to get into the building. And they have your name on a list. So, if there is still a lockdown tomorrow morning, it shouldn't be a problem.

(Whereupon, the above-entitled matter went off the record at 4:35 p.m.)

D

F

N

C E R T I F I C A T E

This is to certify that the foregoing transcript

In the matter of: Neurology Phase II
Steering Committee

Before: NQF

Date: 10-03-12

Place: Washington, DC

was duly recorded and accurately transcribed under my direction; further, that said transcript is a true and accurate record of the proceedings.

Neal R Gross

Court Reporter