NATIONAL QUALITY FORUM

+ + + + +

NEUROLOGY STANDING COMMITTEE MEETING

+ + + + +

MONDAY
APRIL 4, 2016

+ + + + +

The Neurology Standing Committee met at the National Quality Forum, 9th Floor Conference Room, 1030 15th Street, N.W., Washington, D.C., at 8:30 a.m., David Knowlton and David Tirschwell, Co-Chairs, presiding.

PRESENT:
DAVID KNOWLTON, MA, Co-Chair
DAVID TIRSCHWELL, MD, MSc, Co-Chair
DAVID ANDREWS, Patient Advisor, Georgia Regents Medical Center
JOCELYN BAUTISTA, MD, Assistant Professor, Neurology Staff Physician, Quality Improvement Officer Cleveland Clinic Neurological Institute Epilepsy Center
KETAN BULSARA, MD, Associate Professor, Director of Neuroendovascular and Skull Base Surgery Programs, Neurovascular Section, Yale Department of Neurosurgery
JAMES BURKE, MD, University of Michigan
MICHELLE CAMICIA, MSN, RN, PHN, CRRN, CCM, FAHA, Director of Operations, Kaiser Foundation Rehabilitation Center
VALERIE COTTER, DrNP, AGPCNP-BC, FAANP, Advanced Senior Lecturer, University of Pennsylvania School of Nursing
BRADFORD DICKERSON, MD, MMSC, Associate Professor of Neurology, Director of the Frontotemporal Disorders Unit, Massachusetts General Hospital

DOROTHY EDWARDS, PhD, Director, Collaborative Center for Health Equity, University of Wisconsin Madison School of Medicine and Public Health
REUVEN FERZIGER, MD, Director, US Medical Affairs, Merck and Company
DAVID HACKNEY, MD, Professor of Radiology, Harvard Medical School, Beth Israel Deaconess Medical Center
STEPHEN HUFF, MD, FACEP, Department of Emergency Medicine, University of Virginia Health Sciences Center
CHARLOTTE JONES, MD, PhD, MSPH, Director of Quality, Division of Pediatric Neurology, Nationwide Children's Hospital
MICHAEL KAPLITT, MD, PhD, Associate Professor, Associate Professor with tenure, Will Cornell Medical College
RONALD KOENIG, MD, Medical Director, Anthem, Inc.
LISA LINES, PhD, MPH, Health Services Researcher, RTI International
ALEXANDER RAE-GRANT, MD, Director, Cleveland Clinic Center for Continuing Education
MELODY RYAN, PharmD, MPH, Professor, University of Kentucky College of Pharmacy
PETER SCHMIDT, PhD, Senior Vice President and Chief Mission Officer, National Parkinson Foundation
JANE SULLIVAN, PT, DHS, MS, Associate Professor, Northwestern University
KELLY SULLIVAN, PhD, Assistant Professor, Department of Epidemiology, Georgia Southern University*

ROSS ZAFONTE, DO, Professor and Chairman, Department of Physical and Rehab, Harvard Medical School

NQF STAFF PRESENT:

ANN HAMMERSMITH, General Counsel
WUNMI ISIJOLA, Administrative Director
KAREN JOHNSON, Senior Director
ELISA MUNTHALI, Vice President, Quality Measurement
YETUNDE ALEXANDRA OGUNGBEMI, Project Analyst
ANN PHILLIPS, Project Analyst, Health Information Technology
CHRISTY SKIPPER, Project Manager
MARGARET TERRY, Senior Director
MARCIA WILSON, Senior Vice President, Quality Measurement

ALSO PRESENT:

KAREN KOLBUSZ, The Joint Commission
COLLEEN MCKIERNAN, The Lewin Group
LEE SCHWAMM, MGH/Harvard
NAILA WAHID, The Lewin Group
ANN WATT, The Joint Commission

*** present by teleconference**

A-G-E-N-D-A

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

8:41 a.m.

CO-CHAIR KNOWLTON: Okay. We're going to get started.

I'm Dave Knowlton. I'm one of the co-chairs along with Dr. Tirschwell. We'll introduce ourselves in a minute.

I'm going to hand it over to Peg, who's going to lead you through the original - the initial formalities. I do want to ask that you take your cards and turn them towards us until we learn your names because we'll be calling on you to speak and it would be helpful if we could see your names. That would make it a little easier for us.

So David, if you have anything, or I'll pass it off to Peg.

CO-CHAIR TIRSCHWELL: No, just welcome and thank you all for the time. I know you already put in reviewing these measures. We really appreciate your help.

DR. TERRY: Good morning. My name is

Peg Terry and I am the senior director on this project. I want to welcome everybody - I want to welcome the committee, the co-chairs, the developers, those on the phone to our two-day neurology committee meeting.

And I know many of you have spent some time on our work group calls and have gotten to know each other a little bit and have gotten a better look at the measures. So we'll continue from there.

So with that I'd like to turn it over to the staff here at NQF. Christy?

MS. SKIPPER: Good morning. My name is Christy Skipper and I'm the project manager for this project.

MS. ISIJOLA: Good morning, everyone. My name is Wunmi Isijola. I'm administrative director here at NQF looking forward to the next two days.

MS. OGUNGBEMI: Good morning. Alexandra Ogungbemi. Welcome.

MS. MUNTHALI: Good morning. Elisa

Munthali. I'm vice president for quality measurement. I wanted to welcome everyone and thank you so much for serving on the committee.

DR. TERRY: And now I'd like to turn it over to the co-chairs.

MS. ISIJOLA: Well, just before we get started I just wanted to acknowledge also Karen Johnson. Many of you are familiar with her but she's our chief methodologist here at NQF so Karen, do you want to introduce yourself really quickly?

MS. JOHNSON: Hi, I'm Karen. I remember several of your faces from the last time around when I got to sit in Peg's seat. So welcome.

DR. TERRY: And so with that, I'd like to turn it over to the co-chair.

CO-CHAIR TIRSCHWELL: Trying to figure out what's next here but I'm thinking we should go around and everybody introduce themselves really briefly. Is that okay? Or Ann, do you want to do that?

(Off mic comments.)

CO-CHAIR TIRSCHWELL: Well, we already thanked everybody. I don't have anything to add at this point.

DR. TERRY: Great. Okay.

MS. HAMMERSMITH: Hi, everyone. I'm Ann Hammersmith, NQF's general counsel. As your co-chairs just noted, I'll lead you through the disclosures of interest.

If any of you have been in our committees before you're used to this. I don't think my mic is working. It's glowing red.

MS. ISIJOLA: Just to note, we can only have no more than three mics on. So if you're not speaking please turn it off. Thanks.

MS. HAMMERSMITH: Okay. It's not working. Anyway, on to disclosures of interest. As I was saying, if any of you have served on our committees before you know the drill. I'll go over it quickly before we go around the table.

As I said, we'll combine disclosures and introductions because it's a little bit

quicker and we want you to be able to get on with the work we'll be doing today.

So if you recall, you received a disclosure of interest form that asks you a lot of information about your professional activities.

We review those as part of seating the committee. But we do like to do oral disclosures of interest at the first public meeting of a given committee.

The reason we do that is because we want the process to be open and transparent for the public and for all of you and we want you to know where each other is coming from.

So we are interested in your disclosure of items that are relevant to the work before the committee only if they are relevant to the work before the committee.

So if you invented the heart transplant that's wonderful. But it doesn't have anything to do with neurology usually so we don't really want to hear about that.

So please don't summarize your resume. We're particularly interested in grants, research, consulting or significant speaking engagements but only if it has to do with the work before the committee.

So we'll go around the table, tell us who you are, who you're with and if you have anything you want to disclose.

Just a quick reminder, you do serve as an individual on the committee. You don't represent your employer. You don't represent anyone who may have nominated you for service on the committee.

So we'll start with the co-chairs.

CO-CHAIR KNOWLTON: I'm Dave Knowlton. I'm retired - I just retired from being the president and chief executive officer of the New Jersey Health Care Quality Institute.

I've served on this committee with David before and I don't believe - I'm the one y you hardly hear from because I've been in an RV traveling around the country for the past six

months so learning that there aren't cell signals everywhere.

I don't have anything really to disclose other than the fact that I serve as the chairman of the hospital safety score --

CO-CHAIR TIRSCHWELL: David Tirschwell. I'm a stroke neurologist. I work at Harbor View Medical Center, which is part of the University of Washington and Seattle.

As far as disclosures go, I guess I represent or was nominated by Harbor View and I think also the American Stroke Association. And so I think, because I'm on an American Stroke Association committee about quality measures I'm going to be recusing myself from voting during one of the stroke association measures later. That's it.

MEMBER SCHMIDT: I'm Peter Schmidt from the National Parkinson Foundation. I was on this committee in 2013 for the second half of the Parkinson's measures and I have nothing to disclose.

MEMBER J. SULLIVAN: My name is Jane Sullivan. I'm a physical therapist. I'm on the faculty at Northwestern University in Chicago and I am working on a grant for the American Physical Therapy Association on core set of outcome measurement and rehabilitation. But I have nothing to disclose.

MEMBER COTTER: Good morning. I'm Valerie Cotter. I'm an adult gerontology primary care nurse practitioner and I'm on the faculty at the Penn School of Nursing and I have no disclosures.

MEMBER ANDREWS: Good morning. I'm David Andrews. I'm a patient advisor at what's now Augusta University. Some of you probably know it as either any one of several names. It's changed its name several times recently - Medical College of Georgia, Georgia Regents University, among others.

We think that we may have a name we're keeping for a while and I have nothing to disclose.

MEMBER DICKERSON: Hi, I'm Brad Dickerson, a neurologist working in dementia from Massachusetts General Hospital and Harvard Medical School.

I receive grants from the NIH and am on the board of directors in the Alzheimer's Association and other than that have no disclosures.

MEMBER HUFF: Hello. I'm Steve Huff. I'm an emergency - thank you - hello, I'm Steve Huff, an emergency physician and neurologist at the University of Virginia.

I'm here at the request of the American College of Emergency Physicians. I have no relevant research grants or support to the matters of this committee.

Something not on my disclosure sheet - I was on a speaker's bureau for a pharmaceutical company - I think that's been 20 years ago.

MEMBER JONES: Charlotte Jones. I'm a pediatric neurologist - I'm Charlotte Jones. I'm a pediatric neurologist. Okay. I'm still

Charlotte Jones. I'm a pediatric neurologist at Nationwide Children's Hospital and I have two disclosures that I think are important to make.

One of them is is that I am the pediatric representative for the AAN specialty group to the Axon Registry, which is one of - which is the registry that the AAN is identifying in a couple of measures as being used to monitor.

I was not involved with the measure development but I do work - I do have that role with the AAN. And I was on their urgent measure development and none of those measures are being brought forth today.

But those potential interactions with the AAN plus just in full disclosure one of their representatives is three doors down from me, Dr. Patel, who's here as the AAN measure developer. But we have not discussed it and he knows I'm going to say what I want to say anyway.

MEMBER CAMICIA: Good morning. I'm Michelle Camicia. I'm the director for Kaiser Permanente's Northern California Inpatient

Rehabilitation Hospital.

I am past president of the Association of Rehab Nurses who I'm here representing, or not. And I'm also a Ph.D. student at the University of California and doing the psychometric testing of a new instrument not related - nothing to disclose.

MEMBER BURKE: I'm Jim Burke. I'm a stroke neurologist at the University of Michigan and Ann Arbor no financial disclosures.

MEMBER KAPLITT: I'm Mike Kaplitt, a neurosurgeon at Wild Cornell Medical College in New York and I have no disclosures.

MEMBER HACKNEY: I'm David Hackney, a neuro radiologist at Beth Israel Deaconess Medical Center and I have no disclosures.

MEMBER ZAFONTE: I'm Ross Zafonte. I'm at Spalding and Mass General and my disclosures are I believe I was nominated by the American Academy of Physical Medicine and Rehabilitation and in the past I was co-PI of an R24 that looked at training young people related

to understanding outcomes, none of which are being evaluated today.

MEMBER EDWARDS: I'm Dorothy Edwards. I'm a psychologist, professor and chair of the Department of Kinesiology and I'm also a professor of medicine at the University of Wisconsin at Madison.

I am part of the NINDS-funded stroke net. I'm the outcomes person for actually the Georgetown group here in Washington, D.C. and I was nominated by the American Occupational Therapy Association.

MEMBER LINES: Hi, I'm Lisa Lines. I am at RTI International. I'm a health services researcher there. I'm also on faculty at University of Massachusetts Medical School and I have nothing to disclose.

MEMBER RYAN: Hello, I'm Melody Ryan. I'm on the faculty at the University of Kentucky College of Pharmacy and I also have nothing to disclose.

MEMBER BULSARA: I am Ketan Bulsara.

I'm a neuro surgeon out of the Yale New Haven system and nothing to disclose.

MEMBER KOENIG: Ron Koenig. I'm a neurologist in my former life. Now I'm the medical director at Anthem Blue Cross Blue Shield of Georgia -- from David and I'm actually a representative for the insurance industry working with the American Academy of Neurology to set up quality measures but have not done so and as such have nothing to disclose.

MEMBER BAUTISTA: I'm Jocelyn Bautista. I'm an epilepsy neurologist at the Cleveland Clinic. I've participated with the American Academy of Neurology and the American Epilepsy Society to craft clinical guidelines but nothing that pertains to the measures we're discussing.

MEMBER RAE-GRANT: I'm Alex Rae-Grant. I'm yet another neurologist from the Cleveland Clinic. I was recommended by the American Academy of Neurology.

I co-chaired a quality measures

committee in multiple sclerosis and none of those measures are being reviewed. I have no other conflicts.

MS. HAMMERSMITH: Is Kelly Sullivan on the line?

MEMBER K. SULLIVAN: Hi, I am. I'm Kelly Sullivan. I'm an epidemiologist at Georgia Southern University in the College of Public Health with trained specialization in neuro epidemiology and my only potential disclosure is that I'm on the American Academy of Neurology guideline development and dissemination committee and then also was a member of the dementia quality measure development group with the American Academy of Neurology.

MS. HAMMERSMITH: Okay. Thank you. Do any of you have any questions of me or anything you'd like to discuss with each other based on the disclosures this morning?

Okay. One - well, actually two more reminders before I leave. One is I just want to reinforce that you do sit as individuals because

you are experts.

So you are here to give us your opinion but you don't represent anybody else. Also want to ask you to keep in mind that the conflict of interest process can only work if everyone participates in it and is vigilant.

So if you're in the meeting and you think that you have a conflict or you're not sure - you think you might - or if you think a fellow committee member has a conflict or if you think someone is behaving in a very biased way we ask you to speak up during the meeting. We don't want you to sit there and then much later say well, I think I may have had a conflict. We want you to tell us now.

So you're always welcome to speak up openly in the meeting. If you'd rather not do that you can approach your co-chairs who will talk to NQF staff or you can go directly to NQF staff.

Any questions about that? Okay. Thank you.

MS. ISIJOLA: Okay. With that being said and I believe we have one committee member who isn't here yet but once they do come we'll make sure that they disclose any potential conflicts. Thank you, Ann.

So before we get started - there we go - so before we get started we wanted to provide an orientation or more so an introduction of the neurology project.

A lot of this information we've gone over during our orientation. But before we dive into the measures we wanted to just provide an overview of our current portfolio.

So to date we have about 12 endorsed measures that have been reviewed in previous phases. Most of them are stroke measures and this is just a listing of those measures.

And when we say maintenance measures, measures that have been endorsed at some point in time.

Your role as committee members today is to determine whether or not these measures

still adhere to our evaluation criteria - whether or not they're still scientifically sound and if in fact you feel that these measures would provide a robust addition to our current portfolio.

So with the inclusion of those 12 measures during our call for measures, which is at a point in time where we solicit new measures to be included within the project, we receive 14 new measures.

This is kind of different just because we received many eMeasures and we'll be talking about legacy measures in later slides. But we wanted to showcase these to you because they are similar in nature to the previous measures.

Additionally, we have received some outcome measures as well as some hybrid measures.

So as your role as committee members we ask you to take ownership of the portfolio - many of the measures you're looking at today - to determine whether or not they add any value, whether or not they are repetitive or competing

with previous measures that are included in our portfolio.

But during the next two days you'll be able to determine whether or not they suffice.

So previously in previous projects we've still continued to see gap areas where you're looking at the treatment and assessment of different diagnoses of areas such as Parkinson's disease, multiple sclerosis, Huntington disease and the likes.

So as you look at these measures it's also an opportunity for you to signal to developers that we need measures that really address some of these subtopic areas.

So with that being said, that was a brief overview. I'll turn it over to Peg to give us more of a highlight introduction of how we will be evaluating measures over the next two days. Peg?

DR. TERRY: Thank you, Wunmi.

What I wanted to do here is really go over what's new to the process for those who have

been on this before - this committee before.

And so this year for the first time we did something called preliminary analysis and many of you have seen that. Our developers have seen that. The staff has really taken time to review it.

We've had two of our senior directors review all of the measures so you know. I just want to say this is really just a guidance. It is just our review. It is just that.

So I think it's important to understand that. The other thing that's new this year is that we have committee pre-evaluation comments and if you've been looking at the measures recently and throughout the time you can see the many comments that people have put into that part of the measure as we have it.

The other thing that's new is we also have pre-meeting public and member comments and that is also there for people to see.

And so this is to create a more transparent ability to really look at these

measures prior to what we're going to talk about today.

And the last thing is we have evidence and testing attachments which, again, for those who have been looking at the measure you have seen that.

So next I just want to go through something we've done before but we've done this on our Q and A calls - our quality - so and this is to just give you a sense of how NQF is looking at measures today.

So for new measures, and this speaks to evidence, we look at the quantity, quality and consistency and, as many of you know, we have a booklet that we at NQF use and that provides guidance on what is quantity, quality and consistency. So all staff can go through that as well as the committee.

We have - this really helps establish a link for process measures with outcomes and so you'll see that. This is for new measures.

Evidence is a must pass. This year we

have a little bit for maintenance measures and we will - we will be talking about a number of maintenance measures today.

Today there is a decreased emphasis on maintenance measures or at least on the evidence part of maintenance measure and requires that the measure developer to attest that the evidence is unchanged.

If the evidence is unchanged from the last evaluation, the standing committee - this evidence will be, you know, we don't need to spend a lot of time on this particular area.

If change is in evidence then the committee will evaluate as a new measure under evidence.

The second part of the importance to measure in report is gap analysis and this is really - speaks to the opportunity for improvement, variation, quality of care across providers and as well as disparities.

There is an increased emphasis on this and I want to mention also that gap or

opportunity for improvement is also a must pass.

Next. Okay. So under the scientific acceptability we have both reliability and validity for new measures. The measure specifications are precise with all information needed to implement the measure.

Under the maintenance if there's no difference required we do require an updated specification. For reliability and validity including risk adjustment, in particular reliability and validity, there is decreased emphasis.

So if prior testing was adequate there is really no need for additional testing at maintenance with certain exceptions including change in data source, level of analysis or setting. Must address the questions for the - what we call the SDS trial - sociodemographic trials period that we have going on right now.

Next. And so for feasibility - and by the way, evidence is a must pass. I mean, evidence - I mean scientific acceptability is a

must pass.

Going on to feasibility and usability, under feasibility the measure - how feasible is it using this measure including eMeasures feasibility for - that is for new measures for maintenance measures there is no difference. Implementation issues must be more prominent. Under - and this is under usability and use under new measures use - used in accountability applications and public reporting, usability impact and unintended consequences.

There is an increased emphasis in the area of - under maintenance measures and it's under this new process for usability.

Next. So we wanted to talk a little bit about companion measures because you've seen a number of companion measures this time and companion measures are what we call legacy measures that also have an electronic version and many people have looked at this.

Companion measures here are some of the stroke measures with an e-Measure that we

call companion. We have different numbers for those eMeasures but they are basically the same measure except that they are eMeasures.

And then I wanted to spend a few minutes talking about endorsement with reserve status and this is really important today because actually - so this speaks to whether a measure has failed in the areas of gap and remember, gap is a must pass area.

So if the measure has failed but has passed all the other criteria, we have the option of moving that measure to what we call inactive endorsement with reserve status.

This status applies only to highly credible, reliable and valid measures that have high level of performance and many of you have seen this high level performance, especially under some of the stroke measures.

Inactive endorsement with reserve status retains these measures in the NQF portfolio while also communicating to potential users that the measures no longer address high

leverage areas for accountability purposes.

I just want to mention this - that if the legacy measure fails on importance to measuring gap then the companion measure will not pass.

I think that's very important to understand as we move forward today.

Next. Let me just ask you, are there any questions on what I just said?

CO-CHAIR KNOWLTON: Yeah, I want you to repeat that last point because that's very significant. That's very important that we understand - we are going to be dealing with that today. So you may want to go through that again.

DR. TERRY: So let me just say it again. If the legacy measures pass - these are not the eMeasures. These are the paper measures per se. they are the existing measures. They've been endorsed before.

If their gap has what we call topped out possibly then the companion measure will automatically not pass. So those are the

eMeasures.

CO-CHAIR TIRSCHWELL: Can I just ask that - so there's no specific quantitative criteria for topping out? Is that sort of a gestalt that the committee decides on and how does that happen?

DR. TERRY: I think that at this point in time, and maybe Karen can weigh in here, we have - you know, we have looked at the number of years.

There's no number, if that's what you're asking - we've looked at the number of years that it is at what we consider topped out in the high 90s and basically it is - it is up to the committee to decide and, of course, if there are disparities that are mentioned that is another issue that can be taken into account.

Karen, anything else?

MS. JOHNSON: I think the only thing else that I would add is when you're looking at performance data and you see something that you think might be getting towards that topped out

level, be sure to look and see where the data are coming from.

So is that data representing most of the providers and patients that are included in the measure or is it only a small subset. Because if it's a - if it's a subset there may be something else going out there that that performance rate isn't showing it.

CO-CHAIR KNOWLTON: So if the committee was interesting in keeping measure or felt that the, let's say the electronic companion measure was of interest, even though they were really - felt the other one was meeting gap that's how reserve status would come into play. Am I correct? I'm incorrect?

DR. TERRY: No. You can't keep the - the reserve status - if the paper measure does not pass that goes into reserve status. The e-Measure is not even discussed. It cannot - yeah.

CO-CHAIR KNOWLTON: Does anybody have any questions on that? I'm only going on and on about it because I know we're going to be

discussing it and I'd rather make sure we've got a good solid picture before we move on. Everybody okay?

MS. JOHNSON: And let me add just one more thing. Reserve status is created for previously endorsed measures.

So that's why we're saying if something fails on gap and does actually go into reserve status that's fine because it's a previously endorsed measure.

Anything new - that companion measure that's a new measure is not eligible for reserve status because it is a new measure.

DR. TERRY: Okay. Do you want to go ahead, Christy?

MS. SKIPPER: So good morning again. Before I get started, I want to welcome Dr. Ferziger to the table. So if you could please introduce yourself and share any disclosures of interest.

MEMBER FERZIGER: Sure. Thank you and I'm sorry to be late. It's the difficulty of

actually not commuting and living where my kids couldn't get to school this morning.

But I am a psychiatrist. I am a medical director at Merck Pharmaceutical Company in the CMS area and I have no specific disclosures. Obviously, we work on several of the areas in discussion.

MS. SKIPPER: Thank you. So I'm going to go over the role of the standing committee. So the role of the standing committee is to act as a proxy for the NQF membership and as such we expect and we know that you all have a multitude of experiences, values and opinions and so we're looking forward to that discussion today and that collegial interaction amongst one another and even between measure developers.

Standing committee members or standing committees serve two to three year terms and we will be determining that for you all on tomorrow.

And then also we expect that you all will work with us to achieve the goals of the neurology project by reviewing all the measures

that have been assigned to the portfolio and also evaluating each measure against our criteria and then indicating the extent to which those measures meet the criteria and make recommendations to NQF for endorsement.

And we also ask that you respond to any comments submitted during the review period and respond to any direction from CSAC. And then just in general, as we've said this morning, we expect that you all take ownership of the neurology portfolio.

And just a couple of ground rules for today's meeting. Probably the most important one is that we all speak into the microphone.

The proceedings are being recorded and we're also competing with construction right outside the window. So you may be leaning over to talk, as I am doing. But it's really important that we hear the discussion as we move on through the day.

And I won't read every single ground rule on the slide but if anyone goes off task I'm

sure we'll all bring it back to center.

So our process for the measure discussion today - so we do have measure developers in the room today and as their measure comes up for discussion we will be inviting them to the table.

There are two seats seated to my left and so they will start out by introducing the measure for two to three minutes and then the committee - or then the discussion will turn to the lead discussants for the measures.

And so we'll turn it to you all to begin the discussion of the measure. Inside your packet you should have a measure discussion script to help you all along in discussing the measure in the order presented there.

If you don't have a measure discussion script please let us know. And so the lead discussants provide a summary of any pre-meeting evaluation comments and emphasize any areas of concern or differences of opinion.

The developers will still be at the

table to answer any questions that you all may have and also if there's any misinformation that has been shared about the measure they will be able to step in and clear that up for the committee.

And we ask that everyone if you would like to be acknowledged to turn your place card perpendicular to the desk and we will acknowledge you.

And then we'll move on to voting on the measure. So voting on endorsement criteria - we'll start out with importance to measure and report and at this time - at that time we'll be taking two separate votes on the evidence and then one on gaps. Then we'll move to scientific acceptability of measure properties - again, that's two votes. One on reliability and one on validity. Those first two criteria are must pass. If the measure does not pass we will stop discussion of that measure and move to the next measure on the agenda for the morning. So, basically, if the measure fails - either of those

- we'll stop and move on.

Next, we'll move on to feasibility, take a vote on feasibility and one vote on usability and use and Alexandra will give us a more detailed overview when it does come time to vote.

And everyone should have a remote clicker. If you don't have one please let us know.

And then finally, achieving consensus, NQF guidelines state that in order for a measure to pass greater than 60 percent of members must move vote yes and a yes vote is calculated with the sum of high and moderate votes.

A measure does not pass or is not recommended if fewer - if there are fewer than 40 percent yes votes and anything in between that 40 to 60 percent. If that does happen we will address it at that time.

And that is all I have for now. Are there any questions about what I just shared? No? Okay. I will turn it over to Dr. Tirschwell

for our review of the first measure.

CO-CHAIR TIRSCHWELL: According to my agenda, we're a full hour ahead of schedule at this point. So try not to ruin that too badly as the day goes on.

So I have a little script here. I guess so the first measure that we're going to be reviewing as should be available on your agenda and I guess we'll be following along on the screen - you guys will be bringing out the documents for each measure - is 0661, head CT or MRI scan results for acute ischemic stroke or hemorrhagic stroke patients who receive head CT or MRI scan interpretation should be within 45 minutes of ED arrival and the developers are CMS, Mathematica and the Lewin Group.

Is there someone here to introduce the measure or is it by phone?

MS. ISIJOLA: Operator, could you open up the line to see if we have any developers?

CO-CHAIR TIRSCHWELL: Yeah, go ahead. Turn your mic on though please, Steve. Not so

functional.

MEMBER HUFF: So this committee stands for a number of years. If a measure is not approved can it come back in some months for reconsideration or what is that process?

MS. ISIJOLA: So I can jump in and answer that. So if in fact there is additional information that you would like the developer to bring back we can work with the developer to bring that information back and during one of our post calls following this meeting we can address any concerns you have.

CO-CHAIR TIRSCHWELL: And potentially revote and -

MS. ISIJOLA: Correct.

CO-CHAIR TIRSCHWELL: Thank you. Any other process questions as we get started?

And was there a developer available for the 0661 measure that wanted to introduce it?

Okay. Just wait a sec.

MS. ISIJOLA: Operator, could you see if there are any developers from CMS or

Mathematica?

OPERATOR: We have no one from CMS and Mathematica yet.

MS. ISIJOLA: Okay. And I think that maybe due to the time -

CO-CHAIR TIRSCHWELL: Yeah, so is that fair? Should we proceed? Should we wait?

MS. ISIJOLA: Is anyone from the Lewin Group on the phone as well? Operator, could you check?

OPERATOR: No, ma'am. We don't.

MS. ISIJOLA: So what I would say is we could probably have the discussion and if there are information or additional information that we would like to develop or to provide we'll give them that opportunity.

DR. TERRY: We could possibly go to another measure and see if they get on, I mean, to the next measure - to 434. That's a possibility.

Would you - yeah. I'm not sure who --

MS. ISIJOLA: Okay. So what we'll do

is are there developers from the joint commission?

(Off mic comments.)

MS. JOHNSON: Can one of you guys email the other developers to see if they're waiting for the 10/15?

MS. ISIJOLA: Okay. So with that being said, I think we're going to do to the first joint commission measure, Measure 434.

And as Ann Watts mentioned, in the back if there are questions from - for the developer we ask that you hold those and give them an opportunity to address that at that time.

MS. JOHNSON: It might be better just to wait just a second and see if you can get a hold of the other developers via email, just in case they are planning to come on at 10:15.

We're running very early. So maybe we can just pause for a few minutes, go get some more coffee. We'll see if we can get in touch with them just in a couple minutes.

(Whereupon, the above-entitled matter

went off the record at 9:20 a.m. and resumed at 9:32 a.m.)

CO-CHAIR TIRSCHWELL: All right. Now we're really going to start. So as already mentioned, the first measure is 0661 and now we do have some developers in the room.

So you guys are - you have just a few minutes to introduce the measure, appreciate - go ahead.

MS. MCKIERNAN: Absolutely. So thank you for the opportunity to speak today about NQF Number 0661, which is head CT or MRI scan results for acute ischemic - is that better? Should I start over?

So thank you for the opportunity to speak today about NQF Number 0661 which is head CT or MRI scan results for acute ischemic stroke or hemorrhagic stroke patients who received head CT or MRI scan interpretation within 45 minutes of ED arrival.

My name is Colleen McKiernan. I'm a consultant from the wound group and am joined by

my colleague, Naila Wahid, also from them.

On behalf of CMS Mathematica Policy Research and its partner, the Lewin Group, work to maintain NQF Number 0661, a measure originally implemented in the hospital outpatient quality reporting program in 2013 and last reviewed by NQF in 2010.

As you all know, performing prompt brain imaging for patients suspected of acute stroke is a critical component of emergency care for accurate diagnosis and treatment.

Use of a head CT or MRI allows clinicians to differentiate ischemic stroke, hemorrhagic stroke, and mini-strokes. These scans can also help identify candidates for tissue plasminogen activator - or TPA - which is used to treat ischemic stroke patients and is actually contraindicated for treatment for hemorrhagic stroke.

The FDA has approved TPA for use within three hours of symptom onset. Thus, timely imaging following patient arrival in the

ED is essential for rapid identification and treatment of patients have an ischemic stroke.

The denominator for NQF number 0661 captures ischemic - acute ischemic or hemorrhagic stroke patients who arrived in the ED within two hours following symptom onset.

The numerator includes patients from the denominator whose CT or MRI study was interpreted within 45 minutes.

Those under - those under 18, those who expired in the ED, those who left against medical advice or who elected discontinuation of treatment are excluded from the measure.

On behalf of CMS Mathematica and Lewin preformed a series of quantitative and qualitative efforts to assess the measure's evidence-based distribution of performance, scientific acceptability, feasibility and usability.

We look forward to the discussion this morning and are here to answer any questions you may have. Thanks very much.

CO-CHAIR TIRSCHWELL: Thank you. That was perfect. So at this point, we would turn over the floor to the discussants on the committee, Bradford and/or Stephen.

MEMBER DICKERSON: Good morning. Stephen and I - Dr. Huff and I have discussed this ourselves and I'll start the discussion and Dr. Huff will chime in along the way.

So we've heard a summary of the measure. The - this is a maintenance measure that was originally endorsed in 2011 and it's a process measure.

There has been new evidence presented which includes guidelines that subsume several classes of evidence.

So guideline one was three recommendations for patients with acute cerebral ischemic symptoms that have not yet resolved and these guidelines were based on the idea that urgent imaging is recommended to try to treat acute ischemic stroke which has a level of evidence of class 1A and then noncontrast head CT

or MRIs recommended to essentially rule out hemorrhagic stroke and rule in ischemic stroke, also based on strong evidence. The somewhat weaker evidence, I think, here in this guideline is related to the timing. So is 45 minutes a magic number and this is something we'll discuss further. But that has class one level of evidence C.

And then the other guideline, the second point of evidence that was presented is a focused update of current recommendations, again, for treatment of acute stroke and pretty similar.

I think that the overall consensus on the evidence is that there's broad consensus among experts that this is an important goal to be trying to attain.

But I think there is not great empirical research that supports the value of the particular timing, again 45 minutes.

So the idea is that the images should be obtained and interpreted within 45 minutes of the time the patient arrives at the emergency

room and I think, you know, there are questions about where that magic number came from.

But then overall, following the algorithm, the evidence, I think, I moderate. There's overall a high rating for opportunity for improvement. So the idea that we could - and we could talk about where this comes from as well - but the idea that we can improve on our measure I think is a strong opportunity and we can talk more about how we could potentially improve on the measure.

But there's no question, I think, that this - that there's a gap of care and there are opportunities for improvement that we can dig into in a little bit more detail.

So forgive me, I just - this is the first time I've done this and I want to make sure before getting into too much detail that I'm following the process, as you guys want it to be carried out.

CO-CHAIR TIRSCHWELL: I think so. So for - so we sort of have to vote in between

sections. So I guess the idea would sort of be to give your full presentation of information for each of the four sections. Then we'll pause to discuss and then vote, and then we'll move to the next section.

MEMBER DICKERSON: So I think that the first section is the preliminary ratings for opportunity improvement, which the pre-evaluation rated as high and I think - I think I would agree that that is the case. You want to say anything else about -

CO-CHAIR TIRSCHWELL: Steve, any other -

MEMBER HUFF: I don't want to get too picayune here and so what, again, the overall - we're looking at reliability here at the -

MS. ISIJOLA: Right now we're just focusing in on the evidence. Once we have a discussion of the evidence then we will move into voting.

MEMBER HUFF: So I think the summary of the evidence is really that there is an

opportunity for improvement and that there is evidence from CMS that there are disparities and that there are opportunities for improvement in treating acute ischemic stroke and the need to better measure the obtaining of head CT or MRI in acute ischemic stroke patients seen in the emergency room.

CO-CHAIR TIRSCHWELL: So just to clarify a little bit, and I apologize - I wasn't quite specific enough - I guess we're not just voting in between the four sections, we're voting in between subsections?

MS. ISIJOLA: Yes.

CO-CHAIR TIRSCHWELL: So first evidence, then we vote. Then opportunities for improvement, then we vote. So any other comments or discussion about evidence?

MEMBER DICKERSON: Right. So I summarized the evidence which were those two guidelines and I guess we'll pause at that point.

MEMBER HUFF: And thank you for doing the heavy lifting on this. The - we all agree

the intent of this - the intent of this measure is good.

The question is is where does 45 minutes come from. This is level C evidence. People sitting around a table much like this. We don't have any evidence that 44 minutes is good and 46 minutes is bad. We're trying to draw a line in the sands of time.

So there's that. The other issue is it's very surprisingly difficult - I say more difficult in the area of electronic records to determine when a patient arrives in the emergency department.

If you know the patient is coming in in advance, frequently an electronic record is generated and time will pass since - time will pass from when the record is generated to when the patient actually arrives.

Likewise, if providers - doctors, nurses, other people - are with the patient at the bedside they're not doing order entry into that electronic record.

If the patient runs - patient walks into the emergency room - if they come in unannounced there will not be an electronic record in existence.

It takes a few minutes to get that up and running. Providers are with the patient at that time. They've arrived and yet there will be no record of that, and we're left with what we have.

A more reliable marker might be and I believe this has been in other quality measures has been the time from first quarter of their own imaging to interpretation.

That would be a much more granular and a much more reliable measure.

CO-CHAIR TIRSCHWELL: Thanks. Peter?

MEMBER SCHMIDT: So I thought that the 45 minutes came from the fact that they said two - that you're expecting the patient to be in the clinic within two hours and then you've got - FDA label says three hours. So I assume that they had kind of done a gross approximation which, of

course, has some challenges because we don't know.

It's too - you know, the patient is not showing up in two hours, they're showing up in some time and we're hoping it's within two hours and then they said let's give them 15 minutes to administer the drug. That was where I thought that came from, not that that is a justification of it.

MEMBER DICKERSON: Yeah, and I think our point is that so far there's no real empirical evidence to indicate that that's been studied.

But if the FDA says three hours I assume that they've looked at data that says three hours. So whether or not the 45 minutes is based on that two-hour assumption I assume and, you know, I don't know how long it takes to administer TPA but -

MEMBER HUFF: That brings up another discussion, even though TPA is approved for administration within three hours of symptom

onset I would wager that every stroke neurologist here has given it outside that FDA approval - approved time based on other data, the commonly acceptable is four and a half hours. That's not really the main issue with this measure. So that's another point.

But if we're trying to be so granular on this and I don't know that there is really evidence to support that degree of granularity the overall intent is very good. So that's one for this discussion - I think that's probably enough.

CO-CHAIR TIRSCHWELL: Yes, David.

MEMBER HACKNEY: On that same issue, I think the - this may be because they're trying to make it fit a design that is percent of cases that meet criterion X whereas probably what you want is what's the mean time or the distribution of time from arriving at the ED given the vagueness of what that means until you have an interpreted study and then not necessarily have - picking an arbitrary number for cutoff.

But that would require a different design of how you do the measure and all the measures are designed - presented as percent - a numerator, denominator and a percent that meets some criterion. But I agree it would make more sense to say how fast is it?

As for making it from the time you order until the time you have an interpreted study, that kind of, I think, would defeat the goal of the measure which is that whole sequence of events that ends when you have an answer to what the imaging study shows.

You now let everybody off the hook if they simply aren't prompt enough about getting - ordering the study in the first place.

So while it would make it simpler and it would be less ambiguous what the start point was, it probably wouldn't get you to what you actually want, which is fast as possible turnaround.

MEMBER HUFF: I think there's other metrics that address these other issues. So we

all agree there's a sequence of events that takes place and this measure would seem to overreach several different steps in that sequence, and I think by overreaching our steps in the sequence there's going to be some inaccuracy interpreting that.

The other issue that we'll, I think, probably address later on during the discussion is, when is a stroke defined?

So, clinically, somebody comes in, new onset, hemiparesis, aphasia - that's a pretty clear onset. Somebody comes in with some dizziness might not be so clear if they've had a stroke or not - take some time to get that information.

Going further, the question becomes what biomarker are we using to define stroke. If stroke is defined by a MRI showing ischemic changes, perhaps clinically there may be few or little clinical signs to go with that.

And so it's a question of, you know, the interpretation or the measure here says

patients with acute ischemic stroke can almost become a circular reason if that - ischemic stroke is defined by MRI findings and not clinical findings you're going to have much more - a higher failure rate than if it's defined on a clinical basis and that's - I can tell you from the metrics at my institution that happens this patient. You know, what's our definition of who had a stroke. Well, they had a stroke because their MRI was positive. Clinically, it was a very ambiguous setting and yet that drops out the metric. So that's another whole discussion.

CO-CHAIR TIRSCHWELL: Yeah. It turns out from a practical perspective it's all based on what their ICD-9 code is at the time of discharge, which I'm sure does not conform to your much more clinically relevant issue of MRI or what not.

Ketan, did you have a comment?

MEMBER BULSARA: You know, I think when this measure was originally endorsed in 2011 I think it was fantastic - fantastic in the sense

that it actually put a time limit in terms of when the interpretations of studies should be done.

But I think this measure doesn't apply to practice to date and so I'll argue more along the lines of what Steve was saying in the sense that there is new evidence.

There's seven randomized trials that now show that mechanical thrombectomy in a very timely fashion results in improved outcomes.

There are consensus guidelines by many independent societies in which I've been a part of that say patients should be - should go to mechanical thrombectomy within an hour or in 15 minutes from the time of arrival in the emergency room.

So I think what this measure does is I think it - with the 45 minute time frame I think it sends the wrong message and potentially a dangerous message in the sense that the interpretation of the scans need to be done immediately.

You have a patient that comes to the emergency room who's suffering from what we feel is an ischemic stroke. The CT scan needs to be done immediately, like, within a reasonable time frame and the interpretation needs to be done immediately.

It creates an issue on two fronts. The first - the first one it creates an issue on is from the perspective of the patient. If you have to wait 45 minutes to get an interpretation on a CT scan before you administer, let's say, IVTPA before you activate your mechanical thrombectomy team I think you're doing the patient a disservice because it's been shown over and over and over again that the longer time it takes for recannulization the worse the outcome.

I think it does a disservice to the physicians and the treatment team because if you - I mean, you're very qualified. You've been doing this for years. You look at a CT scan. There's no hemorrhage or anything to that extent, in your opinion.

You go ahead and administer, let's say, IV-tPA and at the 45 minute point -- you did this at the 15 minute point. At the 45 minute point, the official read comes back as there's a small hemorrhage and the patient has a catastrophic hemorrhage to follow -- that I think it puts the physicians and the treatment team in a bad position.

So I think this measure is absolutely needed. I think we need a timely interpretation of these radiographic studies. I think the 45 minutes needs to be taken out and I think it needs to be more realistic in terms of what's actually done in practice.

CO-CHAIR TIRSCHWELL: Thank you. David?

MEMBER HACKNEY: Just to clarify, this is not from the time the scan is done until it's interpreted. This is the time from the patient showing up in the ED until the time the scan is interpreted.

So as a radiologist, we often deal

with it was 44 minutes and 30 seconds between the time the person showed up at the ED and the time they got their CT scan. And so we've got 30 seconds to read it.

It's not a time limit. It includes time limits of interpretation but that's not all that's in there, and I think as was previously pointed out, depending on what seems to be going on when the patient shows up, sometimes it's obvious first thing they say this is probably a stroke and at our place, they literally roll the person over to the CT scanner and say this guy's next.

But other times, the evaluation might take a while and that -- while deciding what to do next, before you even order a scan, is included in this measure.

So it is attempting to say everything that leads up to this piece of information you need before you institute therapy counts, whether it's figuring out whether they had a gunshot wound or fell down a stairs or had a stroke,

that's part of that time because if it's a stroke, then you need to get it quickly.

But it doesn't tell you what -- which elements of it are delayed. It says to the hospital. If you have delays then figure out where they are and try to address them.

And I don't know what you do about the fact that there are patients where the presentation is ambiguous.

MEMBER DICKERSON: Excuse me. I think that was part of what Steve and I were wanting us to consider is that if we -- if there's a way to improve this measure, we should make a recommendation about it and one way that we thought of was to identify when the order was submitted because if it's obvious that the person is having a stroke, then the treating team is going to make that order right away.

If it's ambiguous, there's going to be some time for additional work up and that would at least provide some additional information about what the prior probability clinically was

in the treating physician's minds about acute ischemic stroke.

CO-CHAIR TIRSCHWELL: David, did you have a comment?

MEMBER ANDREWS: I appreciate the importance of these issues that the neurologists are talking about. But, as a patient, I would like to have all the right steps done before the treatment started and I'd also like them to be done as rapidly as possible.

So having the evaluation that determines am I having hemorrhagic or an ischemic stroke to determine whether or not you're going to give me TPA, that's an important thing.

If somebody's going to be rushing to try and get it done in a time limit that's unrealistic and then putting me at risk by giving me TPA if I have a hemorrhagic stroke, then that's a concern.

So, you know, my question -- and I guess this is more a question than an observation -- is how much are those people who are actually

delivering the care thinking about oh my god, it's almost 45 minutes -- I have to get it done -- as opposed to thinking about I want to get the best possible care and I don't care if it takes too long?

CO-CHAIR TIRSCHWELL: From a practical perspective, I promise you that probably nobody in the emergency room is -- has got their eye on that 45 minute clock there.

Mostly just trying -- yes, hang on one second, Steve -- to do what's best for you, the patient, every time as quickly and as appropriately as possible.

Ron, Ketan and then Steve again.

MEMBER KOENIG: I don't think you can use when seen by the ER doctor as the criteria. I think it's something that the ER -- because my concern is that if you don't have an adequate protocol in place the persons put in the examining room and half an hour later or 45 minutes later, the doctor shows up and now the clock is running.

No, that's not the answer. The clock starts when the person hits the ER, unless you have a team that is going to start doing evaluations, CT, et cetera, in an ambulance with a mobile service like they have in Germany.

But I think as far as practical U.S. ways when they hit the ER is when the clock starts.

MEMBER BULSARA: You know, David's point is well taken in the sense that there is a lot of issues that lead to potential delays and interpretation of the CT scan.

The issue I have though is that by endorsing a measure that says within 45 minutes, I think what you've done here is you've created extra time -- you've created -- you've given interpretation of a CT scan 45 minutes.

So if your emergency room operation is very, very quick in the sense that your patient gets to the CT scanner within five or ten minutes of arrival or has had a mobile -- I mean, now we're moving to mobile CT scans, right, I mean,

we had some of units floating around. As soon as they walk into the emergency room, potentially, they have a CT scan done.

I think what this measure does is -- I think at least the way I interpret it, in light of today's evidence, is that it gives you 45 minutes to interpret the CT scan that the patient had on the mobile scanner that was done, let's say, a block away, it gives you 45 minutes.

And I think -- I mean, we all strive to provide the best possible care in a rapid fashion and for cerebral ischemic -- for ischemic stroke we know that you have to be rapid. I mean, if you want a reasonable outcome you have to be rapid.

And so I think what this measure potentially allows is allows for unnecessary delay of up to 45 minutes in the interpretation of a very crucial study.

And so I think that's an issue. But David's point is well taken. There are other processes that may limit it but I think you need

to limit the time that the radiologist has to interpret the study.

Has to be done at the -- while the study is done. Sure, I mean, we stand there and we look at it while the study is being done.

MEMBER DICKERSON: So in making a recommendation, sorry, about what to do for the next cycle, would you agree that at least one step in the right direction would be capturing the time from the order entry or the decision that a scan should be obtained to the read itself so that we can quantify that with all the other elements, still being somewhat vaguely assessed?

MEMBER BULSARA: I totally agree with that but I think the other more important -- that's very important but I think the more important question is as the scan is being done, I think the time from this -- the time that this scan is completed to the time of interpretation is an important data element to capture because, I mean, you're making a therapeutic decision in real time based on the information that you're

obtaining. So I think those are -- when is the order put in, when is a scan done, and from the time that the scan is completed -- I mean, CT scan takes less than a minute to complete -- when is that CT scan read, and I think that's the time frame that I think we need to endorse or be more realistic about. I think 45 minutes is unreasonable.

CO-CHAIR TIRSCHWELL: Okay. Bunch of people want to make comments. Alex, Mike, Jim.

MEMBER RAE-GRANT: So it's nice that we have this opportunity to go around a bit before we really get going on these.

But my understanding about our process is that we are not developing de novo -- these measures -- that they come from an external authority and so these come from AHA/ASA guidelines for time lines for things.

So I don't think it's within our purview -- I'm asking -- to modify those guidelines that come to the committee where our job is to decide whether or not we adopt this and

at what level. Is that correct?

CO-CHAIR TIRSCHWELL: You would have to comment on that.

DR. TERRY: Basically, you're asked to review the measure as is and based on the evidence that is provided. Is that --

CO-CHAIR TIRSCHWELL: Yes, I mean -- yes, and so there's a lot of discussion of what might be the optimal practice and the, I guess, Ketan, in some ways you're suggesting there are unintended consequences which we could discuss at a later point in evaluation of all of this about that this somehow allows people to relax and drink some coffee while they're slowly interpreting the CT scan which, of course, is -- you know, I can't imagine any hospital where that's really happening.

You know, from my perspective this measure is more about potentially identifying systems where their percentage might be much lower than somebody else's and they would want to use that information to review their processes.

Maybe they're not identifying them fast enough closest to the front door. You know, I don't -- I guess I would interpret this as giving people license to slow down and I really don't think that people would be interpreting this that way at the hospital level.

But anyway, Mike and Jim.

MEMBER KAPLITT: Yes, I agree. I mean, rather than rewrite this on the fly, the way I approach this is, you know, given what they've shown us, this is supposed to be a minimal quality standard, as you said.

It's not -- this is not the guideline of how you should practice. It's -- this is the minimum quality standard that we feel everybody should try to achieve, right.

And so the way I look at it was is two hours and 45 minutes from last known well -- is the evidence there to suggest that that's -- as an outside limit a meaningful thing to achieve. Hopefully, we do better than that, right.

And that -- you know, while it's true

that the evidence specifically on that point is not provided as well, it's obviously based on the fact that you want to provide some window to treat, given that three hours is the standard for, you know, TPA therapy. There are therapies that you can do faster but that's the minimum. And I think that even though this relates to the next part, which is the performance gap, when you look at the performance gap data it's still so large, which we're about to do in a second, it suggests that while we all would like to achieve faster, the question in my mind from the evidence standpoint is on the assumption that a lot of places are still not achieving this two hour and 45 minutes -- is the evidence there that at a minimum standard this amount of time is valuable for patient quality?

Once we reach the -- once we top out then you could make the argument are there newer standards, I think, that we should achieve.

But I think from the evidence standpoint, the way I look at it is, is there

evidence that this is of sufficient quality that this will do -- that this does something important for patients for the evidence standpoint that we should be maintaining this, you know, recognizing that that's what the last committee had already endorsed.

You know, and so that's the way -- I think that the fact that there's still such a large performance gap we have to say well, does that mean that's because everybody feels that the evidence is pointless or because people are still not achieving this which to me is part and parcel of the evidence in favor of doing this.

MEMBER BURKE: So I think one of the things that's interesting about this measure that hasn't yet been talked about is the -- this is, you know, a diagnostic test.

And so we're talking about sort of a step on the pathway to immediate outcome. If the outcome is giving people TPA, we're now coming up with a dozen complicated logistical problems that are conceivably avoidable if you just say measure

how long it takes to give someone a TPA.

I mean, so the -- if indeed the only causal pathway that we think rapid interpretation is relevant to is administration of TPA, then we now have to figure out is it 45 minutes, who exactly is the denominator, do you have to come up with exceptions for all these other details, when do you start the clock, a million logistics need to be figured out and if the million dollar question is did you get TPA or not or how fast did you give it, then that seems like it should be the question and this seems like it's a strange surrogate that raises a lot of complicated measurement problems that you work around by just skipping to the end point.

I don't know where that fits in the framework. It doesn't really. But it seems like it's relevant.

CO-CHAIR TIRSCHWELL: Yes, Ketan.

MEMBER BULSARA: You know, to Michael's point, I mean, there is great evidence in the sense that -- I mean, we need to

revascularize these patients as fast as possible because we know that if we revascularize them in two minutes we're going to save more of their brain than we are when we revascularize them at 60 minutes.

The point of contention is not that this needs to be done in a rapid fashion or that this need so be done sort of -- like all institutions need to have some sort of minimum standard.

The point of contention is where does this 45 minutes come from. Forty-five minutes is an arbitrary number. It's not consistent with many other societies that have proposed consensus guidelines. Why have we adopted 45 minutes? Why not five minutes? Why not ten minutes? Why 45 minutes? And I think if you have a time like 45 minutes in there I think you have to provide evidence for it and there's no evidence for 45 minutes.

CO-CHAIR TIRSCHWELL: Which guideline gives a different time?

MEMBER BULSARA: So the Society for NeuroInterventional Surgeons has been proposing a lot of recommendations in terms of consensus statements.

And so there's a recent publication that advocates that all organizations should strive for trying to get to revascularization within like an hour and 15 minutes or an hour and 30 minutes.

CO-CHAIR TIRSCHWELL: But that doesn't -- that's not the CT interpretation. Do they give a different time line for CT interpretation specifically?

MEMBER BULSARA: So let's take a step back. So if you're going to activate your mechanical thrombectomy team and if you want them to make growing access within an hour and 15 or hour and 30 minutes there's a lot of steps that occur before then and I think one of the key steps is the interpretation of the CT scan.

You have one hour to activate the team. So most of these strokes happen, let's

say, at 2:00 or 3:00 in the morning. So you have to bring your whole team in. You have about 45 minutes to an hour to bring your team in. And so you have to have --

CO-CHAIR TIRSCHWELL: I understand your point, Ketan, but I'll ask my question again. Do they recommend a different specific time frame for the interpretation of the CT?

MEMBER BULSARA: I'll have to go back to it. I think the recommendation is a rapid real time interpretation of the CT scan.

We don't specify a time.

CO-CHAIR TIRSCHWELL: And the other -- I mean, if you start going down the endovascular route, which is a sort of a different thing, which because the evidence is new, there aren't quality metrics for that are sort of coming through NQF, it actually has to be a CTA, right, not a CT.

So it's really -- it's kind of a different question. If you're moving on to endovascular stuff so --

MEMBER BULSARA: But data -- I mean, some of -- I mean, just out of interest, some of this part of the seven randomized trials, there were some that didn't use any advanced imaging.

They assumed large vessel occlusion based on clinical criteria.

CO-CHAIR TIRSCHWELL: But they all used CTA. I don't think they -- many of them didn't use profusion or diffusion or things like that. But I think virtually all of them used CT angiogram.

MEMBER BULSARA: So I guess the point is, I mean, why -- I mean, in your clinical practice, it's 45 minutes to an hour and we're dealing with 45 minutes interpretation of a CT scan, something that is reasonable.

I mean, do you actually look at the CT scans when you have a stroke patient?

CO-CHAIR TIRSCHWELL: Yes. And I -- to you that this measure doesn't - it's not looking at best practice. It's a way to quantify our total practice in a way that might allow us

to find areas for improvement and I think there's a difference between those two things.

You know, yes, we have residents -- they're in the scanner. The radiology residents in the scanner behind the CT tech and they're interpreting it way before they even write it down. They have to look at their watch so later they can say what time they actually interpreted it.

But, you know, obviously, the resources are quite variable at different hospitals around the country and so you do -- you know, these quality measures are about identifying some reasonable benchmark that will help you identify good practice, potentially identify variance from good practice and would give you the opportunity to focus on some cases where your variance was greatest, maybe the ones where you didn't achieve this measure, because focusing on those cases might help you identify your best targets for improving your processes of care. So there's a few more people -- Melody?

MEMBER RYAN: Thanks. So I think, clearly, when this measure was originally proposed, the idea was to be able to give TPA. Okay. So that would be a three-hour window and that's still the labeling by the FDA.

But what I'm wondering is there are several consensus statements now from the American Stroke Association that suggest 4.5 hours is also reasonable. So do we need 45 minutes or is really like within -- that they arrive within four hours and it's read within 15 minutes or something?

CO-CHAIR TIRSCHWELL: Okay. So you're wondering about specifications it seems like, which might be part of the later discussion.

Steve and then Charlotte and Peter.

MEMBER HUFF: I think we're all in agreement here and we're all in disagreement here.

This is an -- you know, I guess the question is whether to view this measure as an overreaching quality measure or to look at

specific subtypes -- sub-measures within this.

No one's wanting to get emergency physicians off the hook of seeing patients acutely or whoever. There are separate metrics for that. There are separate measures for that. And yes, as an emergency physician, I'm very aware of the clock. I'm very aware that it's very difficult to hit the metrics on our stroke quality measures and sometimes, I have to cut the patient off. I cannot talk to you any longer. I have one minute here to make a decision to pull a trigger on a stroke alert and I can't have you telling me about your TIA last year that left you -- left you weak on your left side. That makes no sense. I've got a minute to make a decision here.

So and -- we're very aware of the metrics. We're trying to hit the metrics. We also want to give good care. So it's almost like we need to define our group. We don't have a good biomarker for this. We're talking about a patient who comes in with a himiplegia and

aphasia. We're talking about a patient who has resolving symptoms. We're talking about a patient who comes in with a headache and nothing else. And so, you know, the CT MRI is our biomarker for this. It's the race to get a reliable biomarker. We have that in cardiac disease with EKGs and enzymes. We don't have that in cerebral vascular diseases. So we're left with our time to the original biomarker. And I think in -- we have issues with both the numerator and the denominator for this measure. It's a real problem.

MEMBER JONES: I think we're all here from ivory towers. There are -- I think we're all here from ivory towers. I practiced for nine years in West Virginia.

There isn't somebody interpreting your CT scan. You're there doing it, and if you are trying to make decisions, if you're trying to perhaps -- there is a neurosurgeon in town who you can call in or you're deciding you have to transport that patient, these time limits, I

think, become destructive when what we're really trying to do is get the patient their treatment in the most -- in the quickest way and moving through.

And when we keep putting time levels on processes, that we're going to grade you on if you do this in a specific time -- if you do this in a specific time, not taking into consideration that the system that you work in may not have these steps, I think the comment that what we really should be focusing on and what NQF has to say is it's the outcome.

And so I think we have to go back to the first question of does the measure assess performance on a health outcome or PRO that impacts the patient, and do we have the evidence that the relationship between the measured health outcome and at least one health care action is identified and supported by the stated rationale. And I -- listening to everyone here, I don't think we have that for the 45 minutes and my understanding is is that's the first thing we're

voting on.

DR. TERRY: I just want to mention that this is a process measure and we do have different requirements based on outcome and process, just to clarify that point. Thank you.

MEMBER SCHMIDT: So one of the things that we struggled with when I was on the panel that it took a little while for us to get clear on this on the panel in 2013 was the -- we have to recognize that not every measure has to target 100 percent as improvement.

So if we think that there are people who will be difficult to get in that window because of presentation, then you have to decide is that going to hurt my clinic because there's going to be some sort of structural bias that I'm going to get more of those than you are and so I'm going to look worse.

And it seems like something where we're getting -- somebody's presenting within two hours of symptom onset, it's going to be fairly randomly distributed the difficult presentations.

And so you don't need to institute an exclusion if the bias is going to be -- you know, if the cases are going to be showing up in a nonbiased fashion.

MEMBER FERZIGER: So I want to ask for some clarification about the point that we're discussing versus the feasibility question.

Because what I'm hearing is that the feasibility, you know, of accomplishing this measure -- I understand feasibility is really about the measure itself -- but its application in different settings will vary. I think what I'm trying to understand is is our job not to determine that there's a benchmark based on medical evidence that ought to be measured and many places may at the current time not be able to meet that benchmark. But that's information that would be very, very valuable. If the benchmark is established, then, you know, all stakeholders can decide how important it is to improve toward the gold standard at their institution. But if we don't measure it, right,

because feasibility of accomplishing the task is poor, you know, then we don't know what to aim for.

So am I understanding this part of the discussion correctly?

CO-CHAIR TIRSCHWELL: I guess I'm -- honestly I'm not 100 percent sure how to answer that question.

MEMBER FERZIGER: Is the question clear or should I --

CO-CHAIR TIRSCHWELL: Could you rephrase it a little?

MEMBER FERZIGER: Sure. Well, what I was hearing is some discussion about whether -- I mean, there are two issues here, one whether the 45 minutes is reasonable from a clinical point of view and two, even if it is, whether it's attainable, given all the other issues involved in care at various institutions.

And it seemed to me that if the second one is part of the discussion now, why would that impact, whether we choose to leave or not.

CO-CHAIR TIRSCHWELL: I think this discussion should be more about whether the 45 minutes is an appropriate thing to do, not whether it's attainable.

MEMBER FERZIGER: So it would seem to me that there was some mixture of those issues.

CO-CHAIR TIRSCHWELL: Okay. Lots of comments. Charlotte, you have another comment and David? Sorry.

MEMBER JONES: Well, just to clarify. The reason I brought up the other institutions who may not be able to meet it is because I don't think that the evidence supports a time, and we have to look at the consequences of putting on a time when we may not have the evidence to support it.

MEMBER HACKNEY: So I agree with the comments about the limitations of this. But I'll just say that from a practical point of view this is useful, and I think I'm restating what Steve had said, that looking at your performance on this helps you identify delays in your system and

it doesn't -- just knowing that number doesn't tell you where the delays are.

But when you dig into your system you can find out where those delays are, and you can try to cut them down and that's where the quality impact of this would lie. Yes, 45 minutes is an arbitrary number.

It's, I think, an attempt and a guess at how long realistically does it take an ED doc to see a patient and decide that the next thing they need is a head CT scan because they might have a stroke.

They can't do that in one second and no, it doesn't take a long time to do or interpret the CT scan once you've gotten that far.

But if you're going to build in all of that, and I think there would be a disadvantage in breaking that down into a tiny bit -- bunch of steps because that lets the system off the hook. This is a system performance measure and it's -- how long does it take when somebody shows up

until you have what we are presuming is the last piece of information you need before you're going to decide to revascularize.

And I think that's a useful thing to measure and I agree, I don't know that there's data that says how long that step -- this step as currently designed should take.

But in terms of usefulness, it helps you find out if you're slower and if you are then you try to speed up, and I think that's the value of it.

CO-CHAIR TIRSCHWELL: Okay. Michael and then Ron and Ketan and then I'm going to say unless you have something that's totally different than what's already been said -- I think we've hit the main themes here, that I would move that we go ahead and vote on this first thing, and I'll note that we are no longer ahead of time. Michael?

MEMBER KAPLITT: Yes. So I completely agree with what you said. I would just clarify again that there is good evidence about time.

There may not be good evidence about this 45 minutes but there is no question stroke is all about time and has good class one, you know, excellent evidence about time.

So we're just arguing right now about the 45 minute number, not about time. A. B, if we all agree around this table that time matters and there's good evidence for most therapies, you know, TPA therapy, three hours, they may have been broadening the window in recent experimental studies but it's three hours -- that's, you know, the way it is, and there may be new therapies with shorter time windows.

But if we agree that three hours has a good basis for it -- there's good class one evidence for three hours with TPA and there's good class one evidence that you need to image people before you provide that, then we're basically narrowing this down to somewhere in that three hours, you have to have evaluated the patient, you have to have done a scan, and that should be the standard of care, at a minimum.

There may be new technologies but that's the minimum.

So yes, you could do a study to say should it be a half an hour, should it be ten minutes, should it be 45 minutes. Then you get into feasibility questions, et cetera.

But I don't think that, you know, suggesting that there's no good evidence about time or there's no good evidence about imaging or whatever, you know, that's kind of the implication of some of this.

I know that might not be the intent, but that's the implication of some of this and there is good evidence of about three hours and there is good evidence about imaging before you provide therapy.

And so now we're kind of parceling it down to just this brief window that we're talking about.

CO-CHAIR TIRSCHWELL: And I would just add to emphasize on that, Mike, that the greater percentage you have that meets the 45 minute the

-- probably the shorter your average time is and that's really what we're shooting for to improve care. So I think there probably is some value.

Ron and then Ketan and then maybe vote.

MEMBER KOENIG: Just as a reference, there was -- I'm sure it's a study but comments made as to when things should be done including within 45 minutes for a scan and that came from the guidelines for early management of patients with acute ischemic stroke, the guideline for health professionals from the American Heart Association, American Stroke Association and that was in Stroke 2013 Volume 44.

I think we just need to have something to look at so we can let EMS know where to go to get the best service and this is, unfortunately, the only way.

MEMBER BULSARA: Just a really quick comment. The data that Ronald mentions is from before the seven randomized studies that showed mechanical thrombectomy as an important player in

terms of revascularizing these patients.

And you had asked about whether in the standards we had established a time and just pulling up the paper it's 15 minutes or less from the patient -- from the time that the patient hits the emergency room to getting a CT scan.

And so these are the ideal standards -- less than 15 minutes -- and I can circulate the article to the group if you'd like.

So the way it's phrased is patient -- I mean, you know if a patient -- the neurology stroke team or the ER stroke team meets the patient in the emergency room. They identify high likelihood of a stroke.

So it's a quick check. They go straight to a CT scan and all of this is done ideally within 15 minutes and the suggestion was that this should be implemented over the course of a couple years and I can circulate the paper to the group.

CO-CHAIR TIRSCHWELL: It's like that's to the start of CT, not to the interpretation.

So it's a bit different.

Ron, do you mind turning off your mic?

MEMBER BULSARA: It's to the -- it's to interpretation. CT scan and interpretation, assuming that interpretation is done by a team that is present there with the patient.

CO-CHAIR TIRSCHWELL: Okay. Last comment, Steve.

MEMBER HUFF: I think we're in agreement this is a non-ideal surrogate marker for summing several processes and I think Jim started off this some time ago that this is a very convoluted thing to some processes.

We want to approve it like this -- this makes sense in this discussion.

I would think going forward though, you know, marker would be -- which percent of TPA-eligible patients get TPA within X minutes. I mean, that's the process we're looking at here.

CO-CHAIR TIRSCHWELL: There is a different measure, I think, that does just that.

MEMBER HUFF: Then perhaps this is

redundant.

CO-CHAIR TIRSCHWELL: Somebody's going to explain to us a little bit about our voting, and then we'll do first -- go faster, I'm really sure --

MS. OGUNGBEMI: Good morning. Everyone should have a little blue remote control. Please let me know if you do not have one.

Once voting is open I will say voting is open and we will make our selections. Please only use your remote control to capture your votes. You can point your remote towards me.

I will capture all of your votes telepathically. There are either two or four options. The two option voting slides are yes or no options, and the four voting or the four option slides are high, moderate, low and insufficient.

Once you press your response, it will display in the little digital screen here. If you press and option that's not available like

number five, because it will never be available, it will show just the small dash instead of an actual number.

If you change your vote while voting is open, only your last response will be captured to avoid multiple votes.

We will give sufficient time to capture votes and then I will say voting is closed and I will read the results of the votes that we've collected.

As Christy mentioned before, 60 percent is required for consensus. So we're going to do a test vote right now.

Pop quiz. What does CDP stand for? Voting is open. You can point your clicker towards me and press your votes.

All right. Voting is closed. I hope this is not a representation of what it actually means or what you think it means. The answer is consensus development process so --

Exactly, yes. Because we met 60 percent on a high or moderate vote, we will

proceed in the criteria.

Okay. So now we will vote seriously.

MS. SKIPPER: And I will be voting on behalf of Kelly Sullivan who was joining us online.

MS. OGUNGBEMI: Okay. We are now voting on evidence for Measure 0661. Voting is open.

The options are high, moderate, low and insufficient.

Voting is closed. The results are 5 percent high, 64 percent moderate, 23 percent low and 9 percent insufficient. Measure 0661 passes on evidence.

CO-CHAIR TIRSCHWELL: All right. We are now moving then to the -- I think it's the gaps section. Can you guys -- are you willing to proceed?

MEMBER DICKERSON: Yes. We --

CO-CHAIR TIRSCHWELL: Take it on -- let's try to keep this one a little more restrained, if we can?

MEMBER DICKERSON: I think this is one place where it looks like this measure is really providing some very valuable information, which basically shows that there is tremendous gap variability in performance and the potential opportunity for improvement, and there are some initial identifications of some disparities with African Americans being less likely than white patients to meet this goal and Hispanic patients being less likely than non-Hispanic, and also female less likely than male to hit this marker.

And patients treated in the facilities with fewer than 50 beds were less likely to have a head CT or MRI scan within 45 minutes of arrival.

The thing that surprised me was that patients treated in major teaching facilities were also less likely, and maybe that's for some of the reasons that we've been discussing today.

So I think -- I would agree that there's a high gap in care and a strong opportunity for improvement here.

CO-CHAIR TIRSCHWELL: Okay. Open for discussion. Seeing none, I suggest we move immediately towards voting.

MS. OGUNGBEMI: We are now voting on Measure 0661 on performance gap. The options are high, moderate, low, insufficient. Voting is open.

Voting is closed. The results are 91 percent high, 9 percent moderate, zero percent low and zero percent insufficient. Measure 0661 passes on performance gap.

CO-CHAIR TIRSCHWELL: All right. Then we're on to scientific acceptability and we start with reliability.

MEMBER DICKERSON: So we could spend as much time on this discussion as we did earlier, I'm sure. But the basic reliability data that have been provided here show a range of .62 to 1.0 with a median of .77 and the statement is that a value of .7 is often regarded as a minimum accepted reliability value. So we would see our measure coming in at .77.

And the summary rating is moderate evidence for reliability and I think I would agree with that?

CO-CHAIR TIRSCHWELL: Any comments, questions? Go ahead and vote then.

MS. OGUNGBEMI: We are now voting on reliability for Measure 0661. The options are high, moderate, low and insufficient. Voting is open.

Voting is closed. The results are 22 percent high, 65 percent moderate, 13 percent low and zero percent insufficient.

Measure 0661 passes on reliability.

CO-CHAIR TIRSCHWELL: So then we're on to validity, correct?

MEMBER DICKERSON: So in validity there was really -- I think this is an area of greater difficulty because, as we discussed already, you know, how do you really -- what's the gold standard, which is often challenging, and I think we could, again, debate for some time on that.

But the validity measure that was provided was really looking at more agreement rate between two different abstractors on 12 critical data elements and so the overall kappa statistic was .52 and I think there were a number of potential threats to validity that were identified and so the summary rating again here is moderate.

And I personally think that we're probably down closer to low to moderate but I could accept moderate, given the challenges in identifying the gold standard here.

CO-CHAIR TIRSCHWELL: Any questions or comments related to validity? Yes, Alex.

MEMBER RAE-GRANT: Just both the reliability and the validity measures, I think -- and those of us who are new to this we need more guidance from methodologists here as to how you guys develop that, how you use it, how we should be interpreting.

I'm not totally clear on what our process is as we go forward on those two

particular parts of the measure -- does that make sense?

CO-CHAIR TIRSCHWELL: Karen, do you mind commenting?

MS. JOHNSON: Sure, let me give you just a really brief overview of testing, particularly for both reliability and validity. NQF allows data element testing or score level testing.

Ideally, we would see both but we don't require both. When we talk about score level testing for reliability, we often see folks do things called signal to noise methodology.

We have some rules of thumb that we provide sometimes. That's where that .7 comes from. It's a rule of thumb to help you think about interpreting what you see. It is only a rule of thumb.

It is not a threshold. So it should not be considered -- you know, if you don't see something at .7 or higher, that doesn't mean that you shouldn't consider it. Again, it is just a

rule of thumb and that's really what NQF does.

We don't specify methodologies, so we don't say you have to do this or the other. We don't specify thresholds, because particularly with reliability and to some extent with validity as well, it kind of depends on the context that you're looking at.

Reliability is a function of the variability that you're seeing both within and between providers, and it also is a function of the number of patients that are included.

So you can see that context matters and that's why it's important for you when you're interpreting testing results to look at the testing sample that was provided.

So, you know, is it -- it doesn't have to be statistically representative but you'd like it to be, you know, big enough to feel pretty comfortable that what you're seeing may be a good signal, if you will, as to what might be happening out in the real world.

For validity, we generally -- for data

element validity we think about comparing -- what you're looking for there is accuracy. So do you feel comfortable that what is being used in the measure is accurate?

A lot of times what developers will do is they will look at something -- if it's a paper measure, they might look at the extraction and then go back and actually compare it to the medical record and see what was extracted, actually what was in the record.

So that's what we see. Sometimes for validity testing, if you are looking at score level validity testing, as opposed to data element validity testing, what you're thinking about there is is there something else that kind of tracks with this measure.

So, for example, if I were looking at the percentage of patients who we got TPA within the three hours -- the limit that you were talking about and you looked later on at functional status of your patient population.

You would think that those two would

track together and that's one way of looking at validity -- at the score level. We would call it construct validity.

So there's a lot of different methods, a lot of different -- statistics that folks could use. That is another reason that we don't lay out what you need to do.

What we try to do in the PA's is tell you whether or not the methodology was an appropriate methodology. If we have rules of thumb, we try to tell you what some of those rules of thumb are.

But, again, you weigh that with what you feel like, you know, is it really hitting that rule of thumb. Maybe it's not quite hitting it but they used a really good sample and I feel pretty comfortable with what we found or maybe not.

I'm not quite sure I answered your question and let me stop there and see if you have additional questions.

MEMBER RAE-GRANT: That's quite a good

start. Thank you very much.

CO-CHAIR TIRSCHWELL: And I'll just add that there are also these flow charts that we've probably all seen at this point, which sort of take you through the various scenarios and the way they've done their reliability and validity testing that I sort of referred to in these green boxes in the documents here.

Okay. So any other comments about -- or questions about validity in this case or should we move to voting? Let's go ahead and move to voting.

MS. OGUNGBEMI: Voting is open for validity of Measure 0661. The options are high, moderate, low and insufficient.

Voting is closed. The results are zero percent high, 77 percent moderate, 9 percent low and 14 percent insufficient. Measure 0661 passes on validity.

CO-CHAIR KNOWLTON: I have a question. I'm noticing that the N is different for the voting. Are some votes not registering? The

previous vote was 23. The N was 23. This is 22.

MS. OGUNGBEMI: Right. Someone did not vote.

CO-CHAIR KNOWLTON: Is that it? What I'm worried is that somebody thinks they voted but it didn't register.

MS. OGUNGBEMI: We can revote if you would like.

CO-CHAIR KNOWLTON: I'm just concerned because I'm noticing that the N is different. You know, sometimes -- I want to be sure.

CO-CHAIR TIRSCHWELL: What is the total N, of all people? Twenty-three? So I guess especially it was a close vote.

CO-CHAIR KNOWLTON: I just wanted --

MS. OGUNGBEMI: And what we can do is let's just revote just so we are capturing everyone's vote. If you can reopen it, just for me.

Voting is open for 0661 on validity. The options are high, moderate, low and insufficient.

The results are zero percent high, 74 percent moderate, 13 percent low and 13 percent insufficient. Measure 0661 passes on validity.

CO-CHAIR TIRSCHWELL: Feasibility?

MEMBER DICKERSON: So feasibility is the extent to which the specifications, including measure logic require data that are readily available or could be captured without undue burden. I just thought I would -- I don't think we've spent as much time talking about that.

So this is available through administrative claims, electronic clinical data, electronic health records, paper and there's actually an electronic data collection tool that's made available from vendors or facilities called the CMS extraction and reporting tool that has some of the data elements required for this measure and it's also feasible to extract from a record by someone other than the person obtaining the original information as was just discussed briefly.

So I think I would agree that this has

moderate feasibility. I think some of the feasibility concerns are related to a lot of the substance of the discussion we had before.

CO-CHAIR TIRSCHWELL: Questions, comments, discussion?

Sorry. Reuven.

MEMBER FERZIGER: Yes, I just wanted to ask how much variance is there across institutions for this particular issue? Is it vastly more feasible in some places than others?

CO-CHAIR TIRSCHWELL: Well, it's -- I think it's feasibility of obtaining the performance measure, not of achieving the performance measure.

MEMBER FERZIGER: Exactly what I mean. That's exactly what I mean. Is it -- there are places that have beautiful IT and quality departments and they are going to be aces at getting this. And there are places, perhaps West Virginia --

CO-CHAIR TIRSCHWELL: Smaller hospitals, we'll just say.

MEMBER FERZIGER: -- that will be very challenged to do this. So it seems like to me like a feasibility of actually collecting the measure might be very different at different institutions.

CO-CHAIR TIRSCHWELL: Yes.

MS. McKIERNAN: So this is Colleen McKiernan. So we -- this measure is among a suite of chart abstractive measures and so your point is well taken. So some facilities might have advanced IT systems whereas others are still using paper potentially.

But because the measures are chart extracted and there's a sample of patient -- of cases that are abstracted for each quarter it really levels the playing field in a way that some of -- some meaningful use measures or other measures that rely on EHR data can require a standardization.

The information is abstracted by a person regardless of the location which the data -- the information is stored.

MEMBER FERZIGER: You probably addressed this, then I thought about it. It seems to me that, you know, when you have -- you have IT systems that automatically are recording times, right, you know, of diagnostic tests and treatments, right, then you have a great deal of accuracy in addition to convenience.

In a place where it's abstracted afterwards, right, you really don't know the accuracy, it seems to me, you know, of the timing that was put down. That would be the lowest priority for many people.

MS. McKIERNAN: So I think that one advantage of the way we have the measure specified currently is that it is not always the time that's documented in the system.

So if you think about the way a radiology report ends up feeding into the system, there might be a delay between when the radiologist entered the information and if a resident entered it in and then the attending had to sign off, so the time that's documented may

not be representative of the actual time and the actual time would be pulled from that note.

So the resident would document the time at which they did the interpretation and then a person would physically have to go in and read that note and find that time.

So you're going to see that, whether it's a nice IT system or a paper record. But, certainly, there are pros and cons of having the EHR measure versus a chart extracted measure.

I think that this is kind of the way it's currently specified now is -- levels that playing field and it does require pulling that information from whatever the physician or whomever wrote down.

MEMBER FERZIGER: I don't want to belabor this but I'm actually concerned about this from a process point of view because when you create a measure where some institutions actually have capabilities of getting accurate data -- is very different than others and there are system incentives actually to meeting the

measure, you know, than it creates some issues about the commitment, you know, of all institutions, you know, to the accuracy, you know, of what they're putting there because there's some pressure on them, therefore some bias, right, to get toward a certain number.

So my concern would be if there is, you know, a significant feasibility variability, we should recognize that and that should be part of how we understand what's collected.

CO-CHAIR TIRSCHWELL: Any other comments? Let's move to voting.

MS. OGUNGBEMI: Voting for Measure 0661 feasibility is open. The options are high, moderate, low and insufficient.

Voting is closed. The responses are 9 percent high, 70 percent moderate, 22 percent low and zero percent insufficient. Measure 0661 passes on feasibility.

MEMBER DICKERSON: Getting to usability and use -- so this is evaluating the extent to which various audiences use or could

use the results for accountability and performance improvement activities, and this publically reported through the CMS HOQR program which is a pay for data quality reporting program implemented by CMS, and the developer reports here that the median rate of head CT or MRI scan for this purpose has increased from 62 percent in 2012 to 71 percent in 2014.

There were some unexpected findings which were -- the wide variation in facility performance doesn't seem unexpected to me but they list that in a report that the median performance is improving.

But they do address the fact that the validity of this measure is challenged by some differences in how facilities and the clinical data abstraction center identified the numerator.

I would assume that would apply to the denominator as well, suggesting that clearer abstraction guidance could improve the validity of the public reported value.

So, you know, I think we would all like to endorse that statement, based on our discussion today if we can.

It's interesting, then, also that the developer states that many facilities are not meeting the minimum case count requirements for public reporting, which are more than ten cases with complete records.

So I think that this is reflective potentially because exclusion criteria may be being applied more variably as well.

And there was no potential for harm reported. So the overall usability described here was concluded to be high, which I would agree with.

CO-CHAIR TIRSCHWELL: Comments or discussion? Let's go ahead and vote.

MS. OGUNGBEMI: We are now voting on usability and use for Measure 0661. Options are high, moderate, low and insufficient. Voting is open.

Voting is now closed. The results are

65 percent high, 30 percent moderate, 4 percent low and zero percent insufficient.

Measure 0661 passes on usability and use.

CO-CHAIR TIRSCHWELL: And then -- discussion of related and competing measures? There's no voting here. We just -- are we going overall? Is that what happens next?

DR. TERRY: Well, I was just going to say we're going to talk about related and competing tomorrow.

CO-CHAIR TIRSCHWELL: Okay. So next we vote on the overall suitability for endorsements. And I guess just -- I mean, it seems like it's passed all the way through and we could still say no?

Doesn't seem quite so algorithmic if that's the case. Right. Okay. Good. Good point. Fair enough.

MS. OGUNGBEMI: Okay. We are now voting for the overall suitability of endorsement for Measure 0661. The options are yes and no.

Voting is open.

Voting is closed. Measure 0661 does meet the NQF criteria for endorsement according to the committee. Thank you.

The responses are 83 percent yes and 17 percent no.

CO-CHAIR TIRSCHWELL: All right. Great. Well, thank you. That was a good first measure. Thank you for tolerating being first.

I think we have a little break, a short break. Maybe we could reconvene -- what do you guys want to say, 11 or five before 11? 11 o'clock and we'll start back in. Thank you very much, everybody.

(Whereupon, the above-entitled matter went off the record at 10:47 a.m. and resumed at 11:00 a.m.)

CO-CHAIR KNOWLTON: Okay. Welcome back to the easy part of our session. We'll zip right through these measures, I'm confident.

We're going to be -- we're going to get an overview from Ann Phillips from the staff

on the issues of eMeasures -- on the criteria and the guidelines and the we will go on to some measures that we're going to measure -- that we're going to evaluate. So Ann?

MS. PHILLIPS: Hi everybody. I'm Ann Phillips and I review eMeasures here at the National Quality Forum.

Most of my work is technical evaluation and getting them ready for your review.

So if we go to, I guess, our introductory slide here we're going to be looking at three different types of eMeasures in the neurology project.

We'll be reviewing candidates for the approval for trial use program legacy measures and one de novo or new measure. That's also a hybrid measure. It's a little more complex than our regular measures.

Some things to keep in mind when considering these measures, especially for trial approval. Trial approval is not an endorsement.

Trial approval is a path to endorsement. We use the trial approval designation for new measures that are innovative that address gaps that need to get out in the field and get more testing data to come back for endorsement.

Hopefully, you will see those measures again while you're still on the standing committee, and you'll continue to review by evaluating the scientific acceptability.

Legacy measures are eMeasures with a currently endorsed claims version that's in use in a federal program, and while the electronically specified version has been implemented in the field, it's been difficult for developers to get complete data on performance right now and we'll accept synthetic testing data for these measures.

You'll review the legacy measures in tandem with their endorsed claims version. It keeps the discussion a little more concise because they are the same measure.

They just have slightly different data elements. And the hybrid measure reviewing is a new measure. It's kind of a unique measure because it's both -- uses claims data to identify the population and data from the electronic medical record to identify the patients -- who's the denominator and numerator.

I think we are ahead a couple of slides. Do you want to go back to the legacy measures for me? Let's talk a little bit about that. Does anybody have any question about the three types of measures, first of all, before we go on to talk about legacy a little more deeply?

Okay. So the legacy measure -- we're going to review this in tandem with the claims measure because it'll just keep the discussion more focused.

Because it makes sense to have one discussion. But keep in mind when you're voting on scientific acceptability for the legacy version of this measure, you can score this area no higher than moderate because we're using the

synthetic tool to support data element validity. So that's important to remember.

You still want to look at feasibility and usability in use and certainly the evidence performance gap, importance to measure in report. Go on to the next one please.

Okay. An approval for a trial use. Does anybody have any questions about the approval for trial use?

Do you all understand it's not endorsement? It's just -- we're just getting good measures out in the field. Testing can be really difficult to get in these -- with these new measures.

So when you review these, you're going to review them against the entire criteria -- the only one part of scientific acceptability and that's Section 2b1, and that that's to determine if the measure specifications are consistent with the evidence. That's a must pass.

So you certainly really want to consider the evidence and performance, the

feasibility and usability in use.

And we -- you know, approval for trial use measures can be supported through a synthetic data set with a testing tool like Bonnie or from a single EHR. They just don't have enough information to satisfy our two EHR minimum for electronic clinical quality measures.

Ideally, these measures will come back to you while you are still here and during your term on the standing committee. They have three years they can be out in the field.

If the measure has not changed, then you'll just talk about the measure on a call and go back and vote on scientific acceptability, and we are starting to see some of our earlier approval for trial use measures when it was still a pilot program before we made it an official program.

We're starting to see these come back. Some of them have come back within six months with enough testing data and the measure isn't significantly different than the measure

originally discussed in the project.

So the committee is able to go ahead and vote on the complete scientific acceptability for endorsement.

Does anybody have any questions about trial approval?

CO-CHAIR TIRSCHWELL: So I'm trying to understand the difference between these legacy measures for which there's no actual data, and approval for trial use, for which there's also no data. How do you get from one to the other?

MS. PHILLIPS: Legacy measures are in use in federal programs. There is a previously endorsed claims version of this measure.

So we can't automatically endorse the electronic version because it uses different data elements. But essentially, it's the same measure. Approval for trial use are new measures. They aren't in use in federal programs. They generally haven't had - they're usually measures that have been conceived relatively recently where legacy measures are

always paired with a claims version that's been previously endorsed.

MS. WATT: Sorry, just a point of technicality. These are not claims-based measures. These are chart-abstracted measures.

MS. PHILLIPS: But we use the claims data to generally identify the population and chart abstraction to go further with the legacies. But these are the electronically specified version. Are there any other questions about approval for trial use?

MEMBER BAUTISTA: Will you be going over Bonnie testing?

MS. PHILLIPS: We'll talk about it a little bit. Do you have any specific questions regarding Bonnie?

MEMBER BAUTISTA: Yeah, what is it?

MS. PHILLIPS: I'm glad you asked. Bonnie is a tool and it's used in measure development and what it allows you to do is create a synthetic data set of patients with various conditions and run the actual measure up

against those patients so you can identify who is in the measure and who is out of the measure, numerator, and denominator. You have the ability to specify ages, genders, give them conditions, set up timing. So the more complex the measure logic the more the complex patients you can develop. But it's not actual patients, so we consider it synthetic.

MEMBER BAUTISTA: And it's not an actual medical record.

MS. PHILLIPS: Absolutely not. It is a synthetic test patient.

CO-CHAIR KNOWLTON: Are there any other questions?

MS. PHILLIPS: Let's go back and talk about the hybrid measure. This is a combination, this is
where you are using the claims data to identify the population, but technically the medical record data. And it's a brand new measure. It's what we call a de novo measure. So the committee should really look at all the information in

testing attachments and feasibility assessment in reviewing scientific acceptability. But this is a measure for full endorsement. It is a new measure. So, three types of measures. You've got, a new measure, a hybrid, approval for trial use and the legacy measures. Are there any other questions about eMeasures? Nope. Okay.

CO-CHAIR KNOWLTON: Thank you Ann. So we are going to move on to our next candidate measures and this 434 venous thromboembolism VTE prophylaxis. And I believe the discussants are going to be Valerie and Peter. Oh, I'm sorry, Joint Commission.

MS. WATT: Hi, my name is Ann Watt, and I'll start. This is one of a series of measures that we include in a measure set. And, I'm going to introduce Karen Kolbusz, my colleague, who is the clinical lead for this measure, and Karen will give you a summary of the measure.

MS. KOLBUSZ: Thanks, Ann. This is stroke one VTE thromboprophylaxis. This is a chart-based process measure. There isn't a

companion electronic measure for this particular measure.

This measure does capture the proportion of ischemic or hemorrhagic stroke patients who received VTE prophylaxis, or who have documentation why no VTE prophylaxis was given on the day of or day after hospital admission.

Excluded populations from the measure include those patients who are less than 18 years of age, patients who have a length of stay less than two days, or patients who have a length of stay greater than 120 days, patients who are documented comfort measures only on the day of or day after hospital arrival, patients who are enrolled in a clinical trial related to stroke, or patients who are admitted for elective carotid intervention.

The measure was originally endorsed in 2012. Basically, the rationale for this measure is that stroke patients are high risk for VTE prophylaxis. Prophylactic therapies are

recommended, since this is one of the highest at-risk hospitalized patient groups. The Class 1 level of evidence A recommendation is for pharmacological prophylaxis.

However, a lesser recommendation, a Class B recommendation, would be for mechanical prophylaxis, in the form of sequential compression devices or intermittent pneumatic compression devices, which may be used for patients when pharmacological therapies are contraindicated.

The measure is widely used, last endorsed in 2012. It is currently in use for Joint Commission hospital accreditation. It's also used in the Joint Commission's Disease Specific Care Stroke Certification programs for primary stroke centers and comprehensive stroke centers. It's used and collected by hospital quality reporting programs since 2013, and is also collected by the Paul Coverdell National Acute Stroke Registry.

CO-CHAIR KNOWLTON: Okay, thank you.

Valerie and Peter, who is presenting?

Peter.

MEMBER SCHMIDT: Yes. So we agreed I would do the play by play and Valerie would provide the color.

So, the -- I'll start with the evidence. So, this has been discussed -- this was passed before and so we are supposed to do an -- as I understand from the introduction, we are supposed to do an abbreviated discussion of the evidence.

The evidence for this measure is clearly good. The only quibble I have here is that all the evidence statements discuss immobilized patients, whereas -- as does the literature, I went back and looked at some of the papers on this, everything specifies immobilized patients, patients with restricted mobility. And yet, the definition of the measure requires -- it does restrict for mobility and it's not an exclusion. And, I was wondering if you wanted to explain that decision.

MS. KOLBUSZ: Patients who did not have restricted mobility could be excluded from the measure, based on a reason for no VTE prophylaxis hospital admission. It would have to be clearly documented since ambulation alone would not be considered a form of VTE prophylaxis.

But if the documentation supported that that's all that was required for that patient, then -- actually, we would not exclude the patient, I correct myself on that one. In this case we would include the case in the numerator population for this particular measure.

MEMBER SCHMIDT: I just want to add, I did actually go down to our stroke center and I asked them about this, and that's what they said, they said that it's an exclusion.

MS. KOLBUSZ: So, we'd give you credit for it, actually.

CO-CHAIR KNOWLTON: Jane.

MEMBER J. SULLIVAN: So, these exclusions came up on several other measures, and I know that it came up on one of our calls. Can

you just clarify the enrollment in any stroke trial and the 120 days, what that -- what the thinking is behind those as exclusions?

MS. KOLBUSZ: The enrollment for 120 days is based basically on a CMS regulation. It has to do with billing practices, since the billing is quarterly and the measure is also collected by hospital and patient quality reporting, we wouldn't want to double bill if the patient was in the hospital for an extended period of time. That's the reason for the 120 day exclusion.

The clinical trial is an exclusion, not only for this measure and this measure set, but many of the core measure sets. The thinking, basically, is that if the patient is enrolled in a clinical trial the usual therapy that would be recommended may not be followed for that patient, so we exclude those cases.

CO-CHAIR KNOWLTON: Peter.

MEMBER SCHMIDT: So, I just want to add I think that the clinical trial exclusion is a

very good one, because there will be a biased distribution of clinical trial enrollees depending on the hospital.

CO-CHAIR KNOWLTON: So, your recommendation on that measure?

MEMBER SCHMIDT: I'm going to go with the recommendation that's in all of your worksheets, it says high. I think that's probably --

MEMBER J. SULLIVAN: And, I would agree with that.

CO-CHAIR KNOWLTON: Any questions from the Committee? And we'll vote on it.

MS. OGUNGBEMI: We are now voting for Measure 0434 on evidence. The options are high, moderate, low and insufficient. Voting is open.

(Voting.)

MS. OGUNGBEMI: Voting is closed. The results are 78 percent high, 22 percent moderate, 0 percent low, and 0 percent insufficient. Measure 0434 passes on evidence.

CO-CHAIR KNOWLTON: Okay, Peter, gap.

MEMBER SCHMIDT: So, on the gap, we see that this had been conducted and that the national aggregate rate on this score has gone from 88 percent to 97 percent, with the mean hospital rate going from 83 to 96 percent. At the 10th percentile, it went from 60 percent to 91 percent, which is a pretty dramatic increase.

And notably, in calendar year 2013 the number of participating hospitals was 264, and in 2014 it was 1,299; so over 1,000 hospitals joined the program without a meaningful decay in the rate of performance.

So, even though that's 1,299 is about 23 percent of all hospitals nationally, it appears that hospitals that joined the program are already conforming to the specification. So, the preliminary rating for opportunity for improvement was low.

CO-CHAIR KNOWLTON: Comments? Valerie.

MEMBER COTTER: I would just like to make a comment in that if clinically we know that there's gaps related to disparities, especially

around ethnic minority persons, it just seems to me that we should continue the measure based on that point.

If the developer does not want to present that data, which I think it would be really interesting, and I think it would be a really important piece to add to this, that that's a reason why I think we should say that we should continue with this measure, even though we have a 97 percent participation rate.

MS. WATT: Hi. This was a --- and thanks for the comment, and this is an issue that came up during the work group calls. And, it's not that we are not interested in disparities at the Joint Commission. It's just that this measure was really not developed for that purpose.

We do collect race and ethnicity, that data element. We just don't slice the data that way ordinarily, and we didn't have it then at the time of the work group. And, actually, we have just within the last 20 minutes got it. We've

been having -- this is true we -- I know you don't want to hear our problems, but we are moving and our data center was having a difficult time getting us numbers.

But, I do have information that says overall if you break down these data in the aggregate, race doesn't seem to make much of a difference.

However, if you look at the individual hospitals, there are a significant proportion of hospitals that do have significant gaps based on the race and ethnicity. And for this particular measure it is 2.1 percent of the hospitals reporting have what our statisticians consider to be a significant gap.

CO-CHAIR KNOWLTON: Any questions on gap? Thoughts? Yes, Jim?

MEMBER BURKE: So, just one question on the temporal trend over time. Do we know how much of this is due to people changing documentation? Are these -- are there specific data on what got you into the numerator? Did you

receive the drug, or did you document the contraindication? Because I think that has some implications for whether or not the persistent gap -- how meaningful that is.

MS. WATT: We have those data. Honestly, I don't have the results and I can't tell you what it is. I'm guessing that there's, you know, a proportion attributable to both.

CO-CHAIR TIRSCHWELL: And so, if I am correct, this vote right here is the one where sort of the topped out issue is in play. And, I don't know whether we need to say that out loud or not. It's, I guess, something to keep in mind, that, you know, as it will be the case with a number of other measures that are very high performing, this is probably the place to consider whether we are there or not.

MEMBER COTTER: And, I think because the disparities data is lacking, that that is a really good argument that this measure needs to be continued in practical use.

CO-CHAIR KNOWLTON: Other comments on

gap? Alex.

MEMBER RAE-GRANT: Yes. Just so to understand the process, so if we do decide this is topped out as a measure, then what's the voting strategy and then what happens to the measure? Just to understand that process.

DR. TERRY: If the measure does not pass under gap, this measure would be eligible to move to the reserve status, and that means it's kept in the NQF portfolio as an endorsed measure, but it's not a measure we go back and review systematically, although we could review it at some point.

CO-CHAIR KNOWLTON: But what would knock it out is if the vote was insufficient, is that correct?

DR. TERRY: Did you say low and insufficient?

CO-CHAIR KNOWLTON: No, I said insufficient.

MS. JOHNSON: So, if the majority of -- a large proportion of folks feel that the

information is insufficient, the measure would go down, and reserve status would not be an option. Okay?

With TJC saying that they have additional information, if you feel that it's -- what Ann has told you verbally is compelling, I think we would probably -- and we could talk about this, but I think we would probably want to continue on. But, we'd really, I think, want to see your data, Ann, and I know it's hot off the press, but we'd probably want to be able to see it.

If it turned out that that was enough to flip you from being topped out to not being topped out, which the disparities could do that, we would, actually, ask Ann to go back and update the submission so that goes on the record, you know, for next time around.

CO-CHAIR KNOWLTON: But, Karen, practically, I'm trying to understand, if we vote and we vote, what knocks it out, just insufficient or low and insufficient?

MS. JOHNSON: Low and insufficient together would knock it out --

CO-CHAIR KNOWLTON: Okay.

MS. JOHNSON: -- but they would have different sequelae.

DR. TERRY: So, for it to move to reserve, all the other criteria would have to pass.

CO-CHAIR KNOWLTON: So, my question goes -- well, go ahead, Peter, then I have a question for him.

MEMBER SCHMIDT: Well, I just want to say, and as I went through this I said that the recommendation from the primary was low. I think based on the fact that we now have some insight that there is additional data that is not on -- in the record, that that should probably be insufficient, because we are aware of something but it's not included.

Is that right? Is that an appropriate interpretation of the intent? But does their verbal presentation of a small amount of what

seems like anecdotal data based on the relatively small ends reported, does that constitute evidence that we can use to adjust our assessment, or should we be basing it on -- it feels like an anecdote.

CO-CHAIR KNOWLTON: Karen?

MS. JOHNSON: Yes, and it's a difficult question, and I want to do that awful thing and pass it on to Elisa.

MS. MUNTHALI: So, what you're voting on is the measure as it is currently specified. But what we can do is work with the Joint Commission during the commenting period, have them update their form, during the post-comment call for the comment period; you can reconsider your vote.

So, that was a good question. The Joint Commission does have information, but we'd like to get more specificity on what that information is. But, just as a reminder, the measure that's in front of you.

CO-CHAIR TIRSCHWELL: And, there's no

option for tabling the vote until the next call or anything like that, so we can review this evidence?

MS. MUNTHALI: In essence, if you voted right now and there's a reconsideration request from the Joint Commission, you'd be revoting during that post-comment call. That's just our standard process.

CO-CHAIR KNOWLTON: Jocelyn.

MEMBER BAUTISTA: So, I thought when Peter was presenting I heard you mention that this only encompasses about 20-some percent of the hospitals nationwide, right? Some 1,200 hospitals, and most of them most likely have Joint Commission stroke certification. So, they are actively trying to improve these metrics.

So, I think the gap then needs to be taken -- needs to take that into account. We are looking at the highest performing hospitals, and it's only 20 percent of the hospitals nationwide who could, potentially, be treating stroke patients.

CO-CHAIR KNOWLTON: Peter?

MEMBER SCHMIDT: So, the interesting point I think is that over 1,000 hospitals joined the program without significantly degrading the average results.

So, as the measure is expanding, we are not seeing evidence of an overall gap, although we have some suggestion that there's some -- that there's a gap in the result findings.

I mean, I kind of feel like if this isn't topped out, what is topped out?

MEMBER BAUTISTA: But, don't you still think the majority of those hospitals have Joint Commission stroke certification, and are --

MEMBER SCHMIDT: I take your statement that they probably do. So --

CO-CHAIR KNOWLTON: Ketan?

MEMBER BULSARA: Sorry. Something I don't understand, so just more for my understanding.

So, clearly, sort of this measure

seems to have tracked the fact that the gap has definitely decreased. So, back in 2010 it was endorsed. So, why -- again, I don't understand this process, I just want to ask it -- why would we not continue to -- even though you are right that there isn't much more gap for improvement, why would we not continue to endorse this measure with the concern that if we decide not to endorse it that we might see an attrition in terms of hospitals reverting back to what they were doing before. So, why would we not continue to endorse it?

CO-CHAIR TIRSCHWELL: You know, I think if you think it's still an important measure, but it's topped out, I think that's the reserve status thing where you are still endorsed, but then I don't actually know what else the difference is, quite honestly, other than we stop measuring it?

DR. TERRY: That is why we have endorsed status -- reserve status I mean, that if it's topped out it can go -- it is still there at

NQF in the portfolio, but it's not a measure that we routinely go back and look at, but we can.

CO-CHAIR KNOWLTON: Is that responsive to you, Ketan?

MEMBER BULSARA: So, if it goes into the reserve, then that means it's no longer actively endorsed, or it's --

MS. JOHNSON: So, it is still endorsed, and we actually have a reserve status policy that we can bring up and let you know. But, the idea would be that it's still there, the NQF still sees that it is a good measure, but the idea that, you know, if you feel that it's topped out that maybe it doesn't really still need to be used in the field, because, you know, there's opportunity costs for collecting data, et cetera.

So, being on reserve status signals to the field that it's still important, it still meets our criteria, but it would not be looked again through our maintenance process. So, it's just kind of there on the shelf if, in fact, performance at some point did deteriorate,

anybody who wants to could come back to say, let's get it off the reserve list, let's pull it back and look at it again to see if the gap actually has decreased.

So, that would be -- and at least -- am I portraying the reserve status, I don't have it right in front of me.

MS. MUNTHALI: You are, and the other piece of that is as a standing committee you have oversight of the neurology portfolio. So, this is a measure that you might want to go back to as experts, knowing that performance and opportunities for improvement may be changing from the time that you last saw it now.

So, it's not like it goes into a closet and we never look at it again. We are just signaling, as Karen said, that this is still a good measure, but based on the information that we have and the data that we have, there are very few opportunities to improve. The performance gaps are not really there.

CO-CHAIR KNOWLTON: Melody.

MEMBER RYAN: So, slightly related -- or pretty much related to this discussion. I'm trying to understand like the practical implications. Does this mean that institutions will no longer monitor this? I would think they would be monitoring it still for Joint Commission.

CO-CHAIR KNOWLTON: I think what this is looking towards -- and NQF staff correct me if I'm wrong -- but I think what happens is there is a wide universe of measures people can measure. So, as performance -- the question is even asked, you know, does this warrant a national performance measure, or have we gotten to a point where it's pretty much adhered to and we can turn our attention to something else where we can see genuine improvements. That's really the issue --

-

(Simultaneous speaking.)

MEMBER RYAN: Sure, and I could see institutions, like, you know, the extractors and whatever, I mean those people cost money. So,

you know, if they don't have to do it or don't want to do it, they can shift to another measure. And, it's clear that having these drives performance. So, I just want to make sure --

CO-CHAIR TIRSCHWELL: Yes. I think, just to clear up the role, the NQF is to endorse or not, put on reserve. The decision about which measures are active or not, via the Joint Commission, is the Joint Commission's decision.

So, it might be a signal to them that they might want to consider refocusing efforts elsewhere, but I don't think it's a mandate to take it off the list in any way.

CO-CHAIR KNOWLTON: Jane?

MEMBER J. SULLIVAN: I guess I have a follow-up question to Melody's. If a measure is on this reserve list, is that the -- is topping out the only reason that a measure would go on the reserve list or how is it publicly communicated why it ends up there? That's one question.

And then my second question is, is

part of our concern to have a parsimonious list? I mean, is that something that we could -- we aren't going to measure everything, and that we are wanting to have a hierarchy of measuring those things for which we get the most bang for, you know, quality improvement.

DR. TERRY: So, it is the only reason to go into the reserve status, topping out.

And, yes, that's the call to have a parsimonious list of measures that are the most effective.

CO-CHAIR KNOWLTON: Peter?

MEMBER SCHMIDT: So, I just want to -- based on the evidence that we've seen, we run the risk of declaring that average is exemplar. And, I think that that's -- that would be a distraction for people who are involved in quality improvement, is to think that a benchmark that we've hit is still a benchmark quality.

And second, you know, the panel and all of us only have so much mind share that we can put to things. And, if we focus on things

where we don't have evidence of a gap, it would wind up being a distraction to future measurements.

And third, it is a very powerful to be able to declare victory on something. If we said, this was endorsed in the past and look at the dramatic improvement, then that is a win. It is not a loss to retire a measure because we've achieved a level of success with it.

CO-CHAIR KNOWLTON: Valerie?

MEMBER COTTER: If however the developer is acknowledging a gap in the evidence, missing data related to racial and ethnic minorities. I have an issue about that. It seems like the data is there, and why is it not being presented?

CO-CHAIR KNOWLTON: I think those are good questions, but we don't have the answer to that, you know? So, we have to vote on what's before us. But, I agree, as this gets brought back, that's a question that we should ask.

It's not --- we don't have the data,

that was Peter's earlier point, we don't have the -- we just have anecdotal. David.

CO-CHAIR TIRSCHWELL: I was just going to say, using my crystal ball it sounds like we will be voting on this one again sometime in the not too distant future. So, even if it does get retired today, the new data on gaps may bring it back into the fold. So, we'll have to see.

CO-CHAIR KNOWLTON: Follow up, Valerie.

MEMBER COTTER: So, where is the push, if you will, or who makes that recommendation that we need this data to move this forward in the future?

CO-CHAIR KNOWLTON: The steering committee doesn't -- the steering committee doesn't develop measures, developers do, but developers are clearly listening to what the steering committee -- believe me, they listen to what you say.

So, my hunch is that we will hear from this again. Ketan?

MEMBER BULSARA: You know, like Peter

and Jane and Valerie pointed out, we get to look at the additional data, but in addition to that, to what Peter was saying in the sense that this is something that we are declaring a victory on. Like with any victory I think you have to understand what it is that is going to safeguard that victory.

So, I think if we retire a measure as important as thromboembolism prevention, I think we have to have a better understanding of what is in place -- not our policy, but what policies have been influenced that will safeguard this very, very important metric. So, I think that would be useful information.

CO-CHAIR KNOWLTON: Charlotte? Comment into your mic.

MEMBER JONES: People have talked about the fact that we have a portfolio to manage. Many of us got 48 hours to review our measures.

We are a limited resource. If we don't retire measures that we feel have been successfully met, and if our -- then for the next

two to three years, and then our colleagues following, we are going to create a situation where we -- as was previously mentioned, we can't do the work that we need to do.

So, I think to say that this is ready for retirement is not saying that we don't think it is valid and important, it's saying we've seen improvement. We don't see a gap right now, it's not going away if someone publishes in three months or a year that there is a bigger gap involving hospitals that don't have joint accreditation, then it's still there for us to pull it back.

But it means that we are not reviewing it a year from now before this meeting and spending the time on it.

CO-CHAIR KNOWLTON: I have a comment on this myself as well, not just to share the measure.

But before I retired I was a quality advocate. And, there's a lot of people going to track these measures and what happens to them.

What the NQF steering committee is asked to do is to set a very high, ongoing bar that people have to keep aspiring to, to get continuous improvement in quality in America, and to retire things that say we've done that one, let's move on.

Nobody -- I hope the advocacy community -- maybe since I've retired there aren't anymore, but hopefully, the advocacy community is going to say -- if this starts to slip, they are going to say, what's going on here? But if the measure has been identified, people are tracking it. What steering committees are asked to do is the hard test of continuing to push that envelope and that's a very important role.

And, you are exactly right, we need to -- and Charlotte's point -- we need to be attending to those variables, and it's a tough task.

So, it's easy to sit back and say, this is wonderful; we don't want to lose it. We

won't lose it, but that's not what we are being asked to do. We are being asked to keep pushing that bar forward, as I see it.

Charlotte, are you still commenting or are you done? Ketan.

MEMBER BULSARA: I'm in total agreement with what Charlotte said. Just, you know, I have these visions of where, you know, if declared victory everything is done, and then you go away.

And so, I think that you are right, we have very limited resources in terms of human hours, but I think it would be important, at least from my perspective, to understand how this has been implemented in policies that are independent of -- how the endorsement by NQF was incorporated into a policy that is being enforced independently by an organization.

I don't think this needs to be reviewed like annually or biannually, or how often that's done, but I do think we have to understand what is it -- what's in place that will ensure that there's not a regression or

attrition in this measure down the road and I'm sure there are policies out there that have always been incorporated, based on the previous endorsements, otherwise we wouldn't see such a high sort of conformity to the measure.

CO-CHAIR KNOWLTON: Alex?

MEMBER RAE-GRANT: You know, I know we are taking the time now to get the process issues behind us, so one thing is one of the issues with the potential gap is the racial-ethnic potential gap we don't know the evidence on.

A question for you guys is do you think about racial-ethnic makeup of your panels, as you constitute them, and should that going forward be something to be considered.

CO-CHAIR KNOWLTON: Well, that's a question for NQF. Go ahead.

CO-CHAIR TIRSCHWELL: So just -- as we move to a vote, I just want to be sure I understand this. Assuming everything else would pass, a vote of low would likely move this towards the reserve list, but a vote of

insufficient would drop it to unendorsed. Is that right? Or, do I have that wrong?

MS. JOHNSON: No, you are almost right.

CO-CHAIR TIRSCHWELL: That's where I usually live.

MS. JOHNSON: It depends. My favorite answer is it depends.

If you land on low, then you'll get a chance to decide if you want to do reserve. And, that will be kind of a gestalt, we won't make you vote on reserve or not, but it's where you want to go.

If it goes insufficient, then technically, it would not pass. Okay?

But, with that, we have heard the Joint Commission has additional data that we would encourage them, and I think they would be willing to bring, so that you could consider it again.

So, with that in mind -- with both of those things in mind, we would still need to go through the rest of the evaluation of this

measure, because even if it went reserve status it has to have passed everything else, right? We have to check that.

If it turns out that they bring something to you later, if you landed on insufficient, we don't want to spend that call doing all the measure. We could go ahead and do everything else but that, and then revisit that again at post-comment.

So, did that answer your question with my it depends?

CO-CHAIR KNOWLTON: So, if the vote was ---

MEMBER FERZIGER: It's actually you that I want to hear from, because I understand and actually strongly agree with the principle you espoused. So, based on that principle, how are you going to vote in this system on this question for this measure?

CO-CHAIR KNOWLTON: Well, let me ask my question first. I still am confused about the insufficient because I thought insufficient says

we would not go to reserve status. That blocks reserve status, and the vote of low leaves that open, is that correct?

MS. JOHNSON: You are correct. So, what we need to do is vote and see where we land --

CO-CHAIR KNOWLTON: So, if we vote --- if you vote for insufficient we are done, we move on.

MS. JOHNSON: We would --

CO-CHAIR KNOWLTON: That's what you waffled on, and I don't think it's -- I don't think that's --

MS. JOHNSON: So, what we will do is if it lands on insufficient, we will continue to discuss this measure anyway.

CO-CHAIR KNOWLTON: Why?

MS. JOHNSON: Because we know that TJC has some additional data that you haven't been able to look very closely at. So, the play off would be, we would have you do that full discussion post comment, or go ahead and do it

now. We have more time today than we do on a phone call later, so we would have you do everything else. It would still go out in our report. If it lands on insufficient, it would go out as not recommended for endorsement, right? But you could reconsider that at post comment.

I do realize this is confusing. It might be more -- it might be easier if you voted first. We'll see where you think you are in terms of is there a gap, because there might be enough information here for you guys to decide that there's not a gap. I mean, I'm not hearing that, but there might be a few here.

CO-CHAIR KNOWLTON: In answer to your question, I don't think there's a gap. I would vote insufficient.

I do believe that we should move this, and if there is new information then I would consider it then. Yes?

MEMBER J. SULLIVAN: I guess I'm a little confused where -- I think what we are being asked to do is vote on the definitions of

low, insufficient, but the conversation kind of seems like we are voting on what should happen.

And I think those may be two different things. The vote would dictate what would happen, but we are asked to vote based on how much of a gap there is, right?

MS. JOHNSON: So, looking at the information in front of you, and hearing what you've heard the Joint Commission telling you about additional information, I think you should vote first on gap.

Once we have your votes, then we will talk about reserve or something else. Okay?

CO-CHAIR KNOWLTON: Peter?

MEMBER SCHMIDT: I just -- can we put up the answer because as far as I understand it, the question is is there evidence of an opportunity for improvement? And there certainly is ample evidence that shows -- that seems to me to show no opportunity for improvement.

So, there's plenty of evidence, it's not insufficient evidence; it's insufficient

opportunity. And we anticipate that the future is going to bring evidence of an opportunity, but we have not seen that in a way that we can critically assess it.

MS. JOHNSON: This is where this becomes your decision on how you want to vote. You could take what's in front of you and land on low, given the percentages in front of you. You could also take Ann's statement that they have additional information that you haven't had a chance to see. And, you could interpret that as not having enough information to make your final decision.

So, it really depends on whether you think you have enough information to make that final call, and that's how you would vote.

CO-CHAIR KNOWLTON: Are you ready for a vote? Go ahead, Charlotte.

MEMBER JONES: Can you restate what we are voting on?

MS. MUNTHALI: I guess the measure as it's currently specified, the submission in front

of you.

CO-CHAIR KNOWLTON: Right. Anything else on discussion? Okay, we are ready for the vote.

MS. OGUNGBEMI: We are now voting on performance gap for measure number 0434. The options are high, moderate, low and insufficient. Voting is open.

(Voting.)

MS. OGUNGBEMI: Results are in. They are 0 percent high, 9 percent moderate, 52 percent low, and 39 percent insufficient.

MS. JOHNSON: So, let me interpret this scenario that I hadn't thought about because -- basically what we have is almost a split between low and insufficient. So, that puts us in a grey zone, quite frankly, as to what we want to do. Okay?

So what we will do is we will put this out in our report as consensus not reached, because we haven't completely -- and Elisa, if I'm saying this wrong tell me -- it hasn't

completely died, but it also hasn't completely gone towards the discussion about reserve status.

So, what we will do is put it out as consensus not reached. Joint Commission will have the ability to bring more in front of you at post-comment call and you will have a chance to rethink and revote.

In the meantime, we will go on to the next criterion, reliability. And let me stop there, see if Marcia and Elisa agree with my interpretation of our regs.

MS. MUNTHALI: Yes, we agree.

(Laughter.)

MEMBER SCHMIDT: So, the next section is reliability. I just want to start out by saying I bet that if you did a yes/no vote, should reserve status be available for this measure, you'd get a 90 to 100 percent yes. So, I think there is a consensus, we are just confused about how we signal that.

MS. JOHNSON: And we can put that in the report as well.

MEMBER SCHMIDT: On reliability, this has been used extensively. We've looked at the data. I think there's a lot of face validity to the fact that this is a reliable measure.

We could probably go through this pretty quickly. The preliminary rating recommendation was moderate, which probably still stands. Do we want to go into more detail about reliability?

CO-CHAIR KNOWLTON: Unless there are questions. Jim?

MEMBER BURKE: One quick one, which is it seems to me like the tricky part here is this documentation of did not receive, or had a reason for refusal or didn't actually -- people who are in the numerator, but did not actually receive the medication. I couldn't find specific reliability testing on that. That seems to me to be the hardest part to determine reliability.

MEMBER SCHMIDT: Yes, I agree with that. It slipped my mind to mention that. I totally agree with you.

CO-CHAIR KNOWLTON: Does the developer have a comment on that, or question?

MS. KOLBUSZ: I think it's how we collect our data being that it's included in the numerator, we aren't breaking it out for reliability because we are giving credit in the numerator. We'd have to go back; we do have patient level data so we have data for the data element.

So, we'd have to go back and specifically look at the breakout of how many included in the numerator wound up in the numerator because of that particular data element, the reason for no VTE prophylaxis. So, I think it's possible to get it, but it's not typically how we report reliability. We look at the exclusions.

CO-CHAIR KNOWLTON: Any other questions or comments on reliability? We move to the vote.

MS. OGUNGBEMI: We are now voting on reliability for measure 0434. The options are high, moderate, low and insufficient. Voting is

open.

(Voting.)

MS. OGUNGBEMI: Voting is closed. The results are 26 percent high, 70 percent moderate, 4 percent low, and 0 percent insufficient. Measure 0434 passes on reliability.

CO-CHAIR KNOWLTON: Okay, Peter.

MEMBER SCHMIDT: So, on validity, the validity testing has been assessed. I assume that there's a typo in the validity testing results, where the decimal point on the P value is in the wrong place, because it says .1 and I believe it's 0.001. If -- given that, I think it's a quite valid measure.

They include here data on frequency exclusions. I noted that I don't see patient mobilized early listed in there, so that's one that I have an interest in, but I'm interested in seeing results on that.

But otherwise, the -- in the preliminary call the assessment was that this was highly valid.

CO-CHAIR KNOWLTON: Comments or questions? Ready for a vote.

MS. OGUNGBEMI: We are now voting on validity for measure 0434. The options are high, moderate, low and insufficient. Voting is open.

(Voting.)

MS. OGUNGBEMI: Voting is closed. The results are 78 percent high, 22 percent moderate, 0 percent low, and 0 percent insufficient. Measure 0434 passes on validity.

CO-CHAIR KNOWLTON: Feasibility. Peter.

MEMBER SCHMIDT: So again, we have quite a bit of face validity on feasibility as it is being implemented in 23 percent of U.S. hospitals. So, I don't see anything that would suggest that this is a difficult or infeasible measure.

The preliminary rating for feasibility was moderate. I would give it a high rating.

CO-CHAIR KNOWLTON: Questions? Comments? Ready for a vote.

MS. OGUNGBEMI: We are now voting for feasibility on measure 0434. The options are high, moderate, low and insufficient. Voting is open.

(Voting.)

MS. OGUNGBEMI: Voting is closed. The results are 74 percent high, 26 percent moderate, 0 percent low and 0 percent insufficient. Measure 0434 passes on feasibility.

CO-CHAIR KNOWLTON: Usability?

MEMBER SCHMIDT: So on usability, this -- in the past year this was used 213,000 times, so it seems pretty usable. And, the recommendation was high.

CO-CHAIR KNOWLTON: Great. Great summary. Any comments? Questions? Let's vote.

MS. OGUNGBEMI: We are now voting on usability and use in measure 0434. The options are high, moderate, low and insufficient. Voting is open.

(Voting.)

MS. OGUNGBEMI: Voting is closed.

Results are 91 percent high, 9 percent moderate, 0 percent low, 0 percent insufficient. Measure 0434 passes on usability.

CO-CHAIR KNOWLTON: Okay. Do you want to go back to -- I hate that word, reserve status?

MS. JOHNSON: Let's reserve reserve status until later. You knew I was going to make you do that Dave.

No, what we will do, since it was consensus not reached I think we should wait to see what Joint Commission has for you.

Depending on what they give you, you may decide moderate, in which case the whole question about reserve status kind of goes away anyway, right? So, it kind of depends on what they give you. So, we will hold off.

CO-CHAIR KNOWLTON: Okay, without objection then we will be moving on to the next measure. Which is --- yes, Peggy, you're trying to get my attention.

DR. TERRY: I just want to let

everybody know, so the next measure, 0435, is really different than the measure we just looked at in that there is a companion eMeasure attached to it. Or a separate measure, but it will -- whatever we vote on will affect the eMeasure. So, I just wanted to make sure you are aware that it's different than the measure we just looked at.

CO-CHAIR KNOWLTON: Everybody okay with that? Any questions? It's what we talked about before, if the primary measure fails then the eMeasure fails.

Okay, this is Dorothy and David. I don't know who is presenting. It looks like David is going for the mic.

MEMBER HACKNEY: I think I'm to present.

So, as discussed, this is the old measure on which the eMeasure is based. So, I think the plan is to go through this in a little bit of detail, and then for the eMeasure only talk about things that are eMeasure specific,

because most of the data comes from this measure.

This is 0435 Discharged on Antithrombotic Therapy. It tries to capture the portion of ischemic stroke patients who were either provided antithrombotic therapy at hospital discharge, or for whom it was documented that such treatment would be inappropriate.

CO-CHAIR KNOWLTON: David, before you go on, I went out of order here. I'd like to hear from the developer first, the Joint Commission wanted to comment. Excuse me for interrupting you, but I'd like to -- if you would like to comment on it, Karen.

MS. KOLBUSZ: Yes, thank you. This is STK-02, which is Discharge on Antithrombotic Therapy. It is a secondary prevention measure. Antithrombotic therapy is found to have benefit in preventing mortality and reducing mortality/ morbidity in ischemic stroke patients.

As the other measures, it was endorsed last in 2012. It's widely used in the Joint Commission certification, and the Hospital

Inpatient Quality Reporting program, and in the Paul Coverdell National Acute Stroke Registry.

The exclusions for this measure are slightly different than the last measure, because it is a discharge measure. This measure also excludes patients who are less than 18 years of age, patients who have a length of stay that are greater than 120 days, patients with comfort measures only documented, patients enrolled in a clinical trial related to stroke. Also patients that would be admitted for elective carotid intervention, as well as several discharge disposition on types of the exclusions, specifically, patients discharged to another hospital or who left against medical advice, as well as those patients who expired, patients who were discharged to home for hospice care, or patients discharged to a healthcare facility for hospice care.

And then, in this case there is a reason data element, which allows exclusion for patients with a documented reason for not

prescribing antithrombotic therapy at discharge.

Now, during the workgroup discussions, there were several comments regarding the high percentage of patients that were reported who were discharged to home with hospice care. I don't know if the Committee would like us to address that, or if we should wait until another point for discussion.

CO-CHAIR TIRSCHWELL: You guys sent us data that we have in front of us, David, are you going to review that briefly when you go through things?

MEMBER HACKNEY: I could, but I think it would make sense for you to do it.

CO-CHAIR TIRSCHWELL: Me, or --

MEMBER HACKNEY: For the -- no, the developer, I'm sorry.

CO-CHAIR TIRSCHWELL: Yes, okay.

MS. KOLBUSZ: Okay, very good. So, the workgroup had pointed out to us that during the workgroup discussions that there were a high percentage of patients that were excluded, based

on that exclusion for patients discharged to home for hospice care.

Patients discharged to home for hospice care are excluded from several of the stroke measures. Basically, this is detected by the discharge disposition data element, which has multiple allowable values, eight in total. Allowable value number two is hospice home, and that's used to exclude those patients who are discharged to home with hospice care.

Looking back and reevaluating the data that was originally recorded, we found that there had been an error in loading data into the data warehouse. Therefore, other data discharge disposition values were being incorrectly counted, along with the allowable value two, which led to us reporting a very high count for this particular measure.

So, in the original data which was resubmitted to this workgroup for consideration, we did have an overall percentage of exclusion reported as 52.2 percent. However, when it was

recalculated to only focus on allowable value two, discharged to home for hospice care, the actual exclusion was 1.29 percent of patients.

MS. SKIPPER: And, you all do not have a copy of that in front of you, but it is available on SharePoint on our home page under General Documents. And, this information was received Friday afternoon, and we posted it this morning, TJC Response under General Documents.

CO-CHAIR KNOWLTON: David, do you want to pick up from here?

MEMBER HACKNEY: So, this is -- as we were saying, this is a measure that's been around for a while. Therefore, it's a maintenance measure.

There's not new evidence about the value of the underlying practice that it's attempting to address, which is discharging appropriate patients on antithrombotic therapy. There's good evidence that this -- doing that reduces subsequent stroke mortality and morbidity.

There are some questions about the -- how should I say, how well it captures the variety of reasons that one -- a patient might be appropriately excluded. So, I will talk about that.

I think the major issue lies in the opportunity for improvement. There is evidence submitted in the measure that --

CO-CHAIR KNOWLTON: Let me interrupt you there. Let's vote on that in order. Let's deal with evidence first.

MEMBER HACKNEY: All right.

CO-CHAIR KNOWLTON: And what your committee recommendation was was high, isn't that correct?

MEMBER HACKNEY: Yes.

CO-CHAIR KNOWLTON: Comments on evidence? Okay, ready for a vote.

MS. OGUNGBEMI: We are now voting on evidence for measure 0435. The options are high, moderate, low and insufficient. Voting is open.

(Voting.)

MS. OGUNGBEMI: Voting is closed. The results are 91 percent high, 9 percent moderate, 0 percent low, and 0 percent insufficient. Measure 0435 passes on evidence.

CO-CHAIR KNOWLTON: Okay, David, gap.

MEMBER HACKNEY: Okay. Opportunities for improvement. There was evidence in the past that the compliance with this measure was not as high as one might hope. The most recent data included in the measure was 98 percent of -- for the 10th percentile rate of compliance, so that essentially, there is a compliance approaches 100 percent. There's very little room for improvement, and given the concerns about whether absolutely everyone who was marked as should have been is appropriately marked as should have been, even that 98 percent might actually be 100 percent of the true group of people who should have been given the antithrombotic.

The other thing that's part of this is evidence of disparities. There are data -- certainly, there are disparities in stroke risk.

There's also summarized in the measure data on the specific goal of antithrombotic therapy at discharge, but it also suggests that these disparities may have declined over time.

The odds ratio indicates a higher performance for some groups than others, but the overall performance is very high for all of the groups reported.

So, I would say there is little opportunity for improvement in overall compliance, and maybe a small amount of potential improvement. Let me rephrase that. I think there's extremely little opportunity for improved compliance overall, and there is maybe a small amount of room for improved performance when you consider this disparity data. But, at least for the groups that were considered there isn't much room for improvement there either.

CO-CHAIR KNOWLTON: Alex?

MEMBER RAE-GRANT: The only point I'd make is the several references I don't think really are germane for concern about

opportunities for improvement in performance. They just tell us that population is different, female outcome, so I'm not sure that's data.

And, I think we've made enough point about racial disparities, but that data should be expected to be developed where potentially available as part of our process. Not that we have to go back each time and say, well, there may be more data, let's get it.

So, going forward, I think it would be reasonable higher expectation if they would provide that where available.

CO-CHAIR KNOWLTON: Okay. Other comments on gap?

Yes, sir.

MEMBER FERZIGER: I have a question about how we think about the disparities gap, is because, you know, if the overall use is extremely high, let's say, you know, it approaches 100 percent, at what point does the disparities gap become non-significant? In other words, do we move on, because the overall number

is excellent, or do we wait because even disparity needs to be addressed?

CO-CHAIR KNOWLTON: I think that is a point, I think that's what Alex was raising, too.

Valerie.

MEMBER COTTER: I think when you are looking at stroke, and you look at the high risk for increased incidence and prevalence among ethnic and racial minorities, that the playing field is not even. And so, while we look at data that does not include ethnic and racial minorities, we can't assume that it's equal.

CO-CHAIR KNOWLTON: I think the issue is that we need data. You can't hypothesize when they say 100 percent of the population are approaching us. That's assuming 100 percent are meeting the standard, if one assumes there are racial minorities in that group. So, you would assume that that's what's happening.

I agree, completely agree, to Alex's point that we should be affirmatively asking for this data. I agree with that point 100 percent.

So, I wouldn't back off that.

But, I don't know that we should just say, no, there's more of a gap here, because we haven't addressed it.

David.

MEMBER HACKNEY: So, I just wanted to say, you know, we talked about, well, if we remove these measures will they, you know, if they expire eventually, will they -- will people stop pursuing these things.

And, what we should be doing is encouraging developers to think about the next step, what is the next step in quality care, and not keep the benchmarks that we set in the past and have achieved.

CO-CHAIR KNOWLTON: Other comments on gap?

David, do you have your --

MEMBER HACKNEY: I just wanted to mention, they do present data on gap. The reason, and maybe what I said wasn't that clear, there was evidence of gap in the past that seems

to be closing, the racial disparity data.

So, they've reported data from a publication from 2010, it doesn't directly say when the data was acquired that had the compliance for White patients at 95 percent, Black patients at 94 percent, Hispanic at 94 percent, and then more recent data that had White at 98, and other race at 98. So, to the extent there existed a gap before, it looks like it's pretty much closed, but they did present the evidence. It wasn't that it wasn't there.

CO-CHAIR KNOWLTON: Anything else on gap? Yes.

MEMBER EDWARDS: Well, I just want to speak to the fact that there appears to be no gap by race in the reporting, but that doesn't mean that there's no gap in secondary prevention, that's a different issue. It's not -- it may be outside the standard.

MEMBER HACKNEY: Right. This doesn't tell us whether overall material is identical, it just says how the performance is on this measure.

CO-CHAIR KNOWLTON: Reuven?

MEMBER FERZIGER: One more question to understand gap. So, we are talking about absolute numbers of patients, we are not talking about institutions or zip codes, right? So, it could still be the case that disparities exist when we compare institutions, or when we compare regions, as opposed to, you know, aggregating the patients. Is that correct? It's a question to the measure developer.

MEMBER HACKNEY: It's certainly true that within that tiny percentage who don't need it there could be some -- that could be driven by some places who do terribly while everyone else does great. That's, certainly, possible, it could be 1.0, at all except a handful of centers that's zero, and you wouldn't know that.

So, if the argument is maybe that keeping it as a reserve status would be a way for them to assess themselves. That's what my suggestion would be, is reserve status.

CO-CHAIR KNOWLTON: Other comments on

gap? Nothing new? Good.

Ready for a vote.

MS. OGUNGBEMI: We are now voting on the performance gap for measure 0435. The options are high, moderate, low and insufficient. Voting is open.

Voting is closed. The results are 0 percent high, 0 percent moderate, 78 percent low, and 22 percent insufficient. The measure passes on performance gaps.

The measure --

MS. OGUNGBEMI: Wait, no, I'm sorry, pardon me.

DR. TERRY: That was a trick question.

MS. OGUNGBEMI: The measure does not pass on performance gap. We'll move on.

DR. TERRY: I just want to make a statement here so it is clear where we are thinking about that, totally clear, as clear as we can make it here.

The measure has failed, this must pass criteria. NQF has the option of granting

inactive endorsement with reserve status for measures that meet all the other criteria except gap. The status applies only to highly credible as well as reliable and valid measures that have high levels of performance due to incorporations due to standardized patient care processes and quality improvement action.

Inactive endorsement with reserve status retains these measures in the NQF portfolio, while also communicating to potential users that the measure no longer addresses high leverage areas for accountability purposes.

The Consensus Standards Approval Committee, CSAC, notes that the default action should be to remove endorsement unless there is a strong justification to continue endorsement.

Does the Committee wish to continue evaluating the measures for possible reserve status? That's the question.

CO-CHAIR KNOWLTON: Well, let me weigh in here on one aspect of this. The issue is, do we, isn't one of the issues that we are

addressing here is keeping this alive so we can consider the eMeasure, isn't that the issue?

MEMBER HACKNEY: What happens if it goes to reserve, does the eMeasure just never come up?

CO-CHAIR KNOWLTON: That's the confusion here.

MEMBER HACKNEY: I would just say there's a logic to saying it doesn't come up, because all the eMeasure is intended to measure the same thing. So, if we think that compliance is so high that there isn't much weight --

CO-CHAIR KNOWLTON: I understand.

MEMBER HACKNEY: -- then I'm not sure there's a point in having much of a discussion of the eMeasure. But, I don't know what the rules are.

CO-CHAIR KNOWLTON: But, if you don't consider -- I agree with you, I just want to be sure the process is clear -- if we don't consider the reserve status of the eMeasure, we don't even look at it. We're done.

MEMBER HACKNEY: So, I guess because the eMeasure passed to reserve status, I guess we're heading to reserve.

MS. JOHNSON: Right, and since this is a new measure to NQF, a new measure that's never been to NQF before, it is not eligible for reserve status.

So, what we were thinking is that you have said that there is very little opportunity for improvement for this measure. That means that the eMeasure also would have very little opportunity for improvement, and we would stop discussion of the eMeasure, if we were talking about the eMeasure right now. We can continue talking about this measure, because it has potential for reserve status. The eMeasure does not have potential to go to reserve status.

MEMBER HACKNEY: So, that makes sense. The problem is, of course, people are going to switch to the eMeasure, right? I mean, if you do this at all. And so, I'm not -- we don't know what to do. Should we go ahead and treat this

one, even though we said that the gap is low, keep going, and then -- you know, I think if we want to vote on whether it goes on to reserve status we have to go through the rest of the criteria to make sure it passes all of those.

So, I guess I would propose that we do that. Then if it does go to reserve, what I'm hearing is that the eMeasure is just off the table. Okay?

Right. And, it seems like going forward if it's on reserve it would really be easier to measure the eMeasure. So, I think maybe there needs to be a little more discussion at the policy level at NQF as to how to think about these things. Maybe it's more of a transformation than a whole new measure.

CO-CHAIR KNOWLTON: Melody, go ahead.

MEMBER RYAN: Okay. This might be silly, but can we approve an eMeasure and then immediately put it on reserve?

CO-CHAIR KNOWLTON: No.

MEMBER RYAN: Okay, just thought I'd

throw it out there.

MEMBER HACKNEY: That beginning is going to fail on gap also. It's guaranteed to fail on gap. This is using the same data.

CO-CHAIR KNOWLTON: My suggestion, and I'm chairing this section, is let's go forward on the measures, on the rest of the measures. And, let's vote on them, so that that question is at least done. And then, we can decide on reserve, but we can discuss it over lunch and come back. Let's get through the criteria, David, if you don't mind.

MEMBER HACKNEY: Okay. So, we are up to reliability, right?

CO-CHAIR KNOWLTON: Right

MEMBER HACKNEY: So my script has a restatement of numerator/denominator, et cetera, but that was already covered.

There is -- the measure does produce good evidence of thorough testing and showing the measure to be reliable, that is, you get the same result when you repeat the test. So, I would

give it high marks for reliability.

CO-CHAIR KNOWLTON: Comments or questions? Okay for a vote?

Go for it.

MS. OGUNGBEMI: We are voting for measure 0435 on reliability. The options are high, moderate, low and insufficient. Voting is open.

Voting is closed. The results are 57 percent high, 43 percent moderate, 0 percent low, and 0 percent insufficient. Measure 0435 passes on reliability.

CO-CHAIR KNOWLTON: Jocelyn?

MEMBER BAUTISTA: Just to clarify something. So, for the reliability algorithm, isn't that that if the testing is only done at the data element level, and not the measure score level, the highest rating possible is moderate?

MS. JOHNSON: That is correct according to our algorithm.

CO-CHAIR TIRSCHWELL: So, are we allowed to vote the seat of our pants outside of

the algorithm, or --

CO-CHAIR KNOWLTON: Yes.

MS. JOHNSON: Absolutely. I think it might be helpful if anybody is willing to share, and sorry, I was kind of thinking about something else when that went through. I see a lot of highs on there, which is a little bit outside our algorithm. So, maybe a couple people who voted high might be willing to share why they felt it was high. That would help us at least understand, as David said, voting outside the algorithm.

CO-CHAIR KNOWLTON: My hunch is they didn't contemplate the algorithms. Your point is well taken, though.

Can we move on to validity?

MEMBER HACKNEY: My program took advantage of that little break to crash. I had to restart. We are back now.

So, validity, basically, was addressed as base validity, that is, it was surveys and focus groups of the hospitals participating in

the pilot test, and they all agreed that it worked. They have a mechanism for ongoing feedback about the performance of the measure, and they didn't get a lot of feedback, but they got described as the specific issues that came up.

The primary one, I think, there were two primary issues. One was this issue that I mentioned, that is, whether the search for documentation of a reason, we are not sending someone out on any thrombotic therapy successfully identified all of the patients for whom it was appropriate not to put them on such therapy. That, the description of the process was a little worrisome, and, of course, there's no real validity possible to know whether a given case was correctly characterized.

And then, the other validity question that I think is more directly handled is, as the number of drugs that one might use keeps enlarging how up to date is the data that is the source table list of acceptable drugs. They say

they do a quarterly review, which sounds like it could be plenty, but I don't know what happens if you scored somewhat as not an inappropriate drug, but, in fact, that is now FDA approved, or it's not FDA approved, but it's accepted practice to use that drug.

In neither of those is there a straightforward of getting validity data for those questions. So, nonetheless, based on what they can present, that is, essentially, expert opinion, it's believed to be highly valid, but I might like to hear if there are neurologists who have an opinion, I'd like to hear those, because now we are getting into a clinical detail, on which I have only a peripheral engagement.

CO-CHAIR KNOWLTON: Any comments?

CO-CHAIR TIRSCHWELL: I'd just comment from a clinical perspective, I would say it's not that ambiguous, and it's probably pretty valid. It's so complicated.

MEMBER HACKNEY: In that case, I would give it a high validity. Everything else about

the validity is good.

CO-CHAIR KNOWLTON: Anything further on validity?

Ready for a vote?

MS. OGUNGBEMI: We are now voting on the validity for measure 0435. The options are high, moderate, low and insufficient. Voting is open.

Voting is closed. The results are 70 percent high, 30 percent moderate, 0 percent low, and 0 percent insufficient. Measure 0435 passes on validity.

CO-CHAIR KNOWLTON: Feasibility, David?

MEMBER HACKNEY: So, I would give it a high. The only issue I brought up for my notes about feasibility was this question of how well does the document in the reason for not using antithrombotic therapy work, and David said he thinks this isn't going to be a problem.

So, that was the only thing that I had any concern about. Otherwise, and that was the one that it was unclear how you would know

whether that was an issue.

They did mention a few people that commented on it, but it's in wide use, and this small number of comments suggested to me that people weren't complaining about that. So, I think it's feasible.

CO-CHAIR KNOWLTON: Comments?

Ready for a vote.

MS. OGUNGBEMI: Okay. We are now voting on feasibility for measure 0435. The results are high, moderate, low and insufficient. Voting is open.

Voting is now closed. The results are 48 percent high, 52 percent moderate, 0 percent low, and 0 percent insufficient. Measure 0435 passes on feasibility.

CO-CHAIR KNOWLTON: Usability and use.

MEMBER HACKNEY: So, usability, it's being widely used by National Stroke Registry, the Joint Commission, and CMS. The data are publicly reported aggregate level, obviously, and there is, as we discussed, evidence of

improvement over time to the point that it may not be useful to keep checking.

It's unclear whether the measure -- to what extent one can attribute the improvement in performance over time to this measure was probably part of it, but, of course, there have been major efforts across the neurology world to improve compliance with guidelines. And, this is one element of that effort.

But, I don't know that we can isolate how much of the improvement is due to this measure. My guess would be not that much, because, as I said, this work getting with the guidelines has been such a big area over the same period of time that the improvement was documented, the quality improvement was documented.

But overall, I'd say it's definitely usable by all of the criteria that we were asked to address. So, I would give it a high usability.

CO-CHAIR KNOWLTON: Comments?

Questions?

Ready to vote on usability and use.

MS. OGUNGBEMI: We are now voting on usability and use for measure 0435. The options are high, moderate, low and insufficient. Voting is open.

Voting is closed. The results are 83 percent high, 17 percent moderate, 0 percent low, and 0 percent insufficient. Measure 0435 passes on usability and use.

CO-CHAIR KNOWLTON: Okay. Done?

MEMBER HACKNEY: Yes, done.

CO-CHAIR KNOWLTON: So, we are not considering 2832, it's an eMeasure, and we don't consider it because the underlying measure did not pass on gap.

And, I spoke to staff and said we'll talk about the reserve status during our break, and we won't do anything on it now.

So, that leads us to --

MEMBER HACKNEY: Let me just interrupt for half a second.

CO-CHAIR KNOWLTON: Sure.

MEMBER HACKNEY: So, the reserve status thing, is that a vote later, or is that -- what's the process for making that decision now about reserve status?

MS. JOHNSON: It should be a vote, and it could happen now.

MEMBER HACKNEY: It could happen now?

MS. JOHNSON: Yes.

MEMBER HACKNEY: Is there a reason to delay?

CO-CHAIR KNOWLTON: I didn't know that the process was clear, that's why I was putting it off. If the process is clear, we can vote upon it. I was --

MEMBER HACKNEY: For data that we thought was going to be given to us, and it --

CO-CHAIR KNOWLTON: Yes.

MEMBER HACKNEY: -- so I think the more data thing might allow this measure to come back in and have another look at it?

MS. JOHNSON: That was the prior

amendment.

MEMBER HACKNEY: That's not this one.

MS. JOHNSON: Yes.

CO-CHAIR KNOWLTON: So it was only because --

MEMBER HACKNEY: It was only because the other one had the insufficient part, not the low part. It was inconclusive as to which of those was the right thing. I see.

So, this one is not coming back with evidence of gap, and the question is, should it be then reserve status. That's the only remaining question. So, I would suggest while it's fresh in our minds we go ahead and vote about it.

CO-CHAIR KNOWLTON: Okay.

CO-CHAIR TIRSCHWELL: Can I just ask what exactly does reserve status mean, when something goes on to reserve status, then what?

MS. JOHNSON: If you'll bear with me just a second, I'm going to pull up our policy and just read to you a little bit. I want to

make sure I get it right.

And, if you want to follow along with me, it's in our Steering Committee guidebook. I'm sorry, it's in our criteria and guidance document, which, hopefully, you have had a chance to look at.

"Endorsement with reserve status retains ...," let me start again, "The purpose of an inactive endorsement with reserve status is to retain endorsement of reliable and valid quality performance measures that have overall high levels of performance, with little variability, so that performance could be monitored as necessary to ensure that performance does not decline. The status would apply only to highly credible, reliable, and valid measures that have high levels of performance due to incorporation into standardized patient care processes and quality improvement actions if the issue for continued endorsement is the opportunity cost associated with continued measurement and high levels of performance, rather than focusing on

areas with known gaps of care. Endorsement with reserve status retains these measures in the NQF portfolio for periodic monitoring, while also communicating with potential users that the measures no longer address high-leverage areas for accountability purposes."

CO-CHAIR TIRSCHWELL: And, can I just comment that this issue about the trade-off being the burden of data collection sort of plays right into this whole eMeasure thing, because, in fact, if you have a good eMeasure it's all of your data is collected as part of your standard process of care. And so, that burden goes away.

So, I mean as the electronic medical record is emerging, and the possibility to gather these quality data become less burdensome, sort of that definition of what constitutes reserve may need to evolve to accommodate this possibility.

MS. JOHNSON: We definitely have these kind of conversations a lot here at NQF, so we will definitely take it back. What we are

hearing from you is you are uncomfortable not putting the eMeasure in reserve status as well. At least that's how I'm thinking about it.

I'm not going to promise that we will change our policy at this point, but I will say that we can discuss it in the future.

MEMBER HACKNEY: Yes, I think the other aspect of that is that this measure, whether -- whatever status it's in, everyone is going to end up doing it by electronic methods. So, we may have a strange situation where we have a reserve status of a technique that no one is going to use anymore, and no approval of what everyone would, actually, use.

We are not going to resolve that in this Committee, but just I think that feeling might have been all the data are the same, the eMeasure is, actually, the more interesting one, because that's what's going to apply in the future, rather than in the past.

But, I guess we are just going to vote on reserve at this point, right?

MS. JOHNSON: And, I will say that the eMeasure, how it functions in terms of its reliability and validity is still unknown, right? So, we can't automatically put an unknown eMeasure into reserve status at this point. So, it's kind of a circular problem.

CO-CHAIR KNOWLTON: Michael.

MEMBER KAPLITT: Can you guys put back that voting slide with the criteria on it, because I had a question on one of the elements in there.

So, how important, in terms of moving to reserve status, is demonstrated improvement, because I don't know how that fits into this reserve question, because to me it doesn't show any improvement. People were doing this equally well five years ago and every year following that. It's not like it's shown improvement over time, and now it's kind of topped out. It looks like it was topped out back in 2010, if I'm understanding the data properly. So, how does that enter into this?

Like I don't know what role that plays? Does that mean that it shouldn't be reserve status if it hasn't shown improvement, because people have been doing this standard of care forever, and so why bother? I don't understand what that line item is for.

MS. JOHNSON: You know, that's a really good question. We may need to relook at that and see what we were thinking.

My understanding of this is that, you know, generally, a measure would not have been endorsed the last time around if there was not at least some flavor that there was opportunity for improvement. And now, we are looking at it again, and we are hearing that there's not.

So, just based on that, what we would assume happens is that, in the time period from the last endorsement to this one there has been at least some improvement, and that would be why that has been there.

CO-CHAIR TIRSCHWELL: If you look at the 10th percentile, Mike, I realize it's not a

lot of improvement, because there wasn't a huge amount of area, but, you know, the 10th percentile has gone up from 94 to 98. So, you know --

MEMBER HACKNEY: And, the disparities have gone down a bit.

CO-CHAIR KNOWLTON: Karen?

MS. JOHNSON: I will also tell you that a reserve status is not new. What is new, actually, is just our written policy of how we are doing this. And, I don't remember if neuro was one of the ones the last time, for some reason I was thinking it was, there was discomfort in various committees of saying, yes, this measure is topped out, and we've already talked about this measure.

It seems that we topped out, but yet, we feel very uncomfortable not endorsing it, right? So, what was happening is, we were -- some committees were saying let's go reserve, others were saying let's not endorse, and it was kind of inconsistent.

So, what we tried to do is make a reserve policy, at least consistent across committees, so that, you know, that feeling of discomfort by signaling to the field that you don't think it's important anymore, we tried to take that discomfort away from you by offering potential reserve status.

MEMBER ANDREWS: My concern is not improvement, it's whether or not the absence of endorsement leads to decline in the use of something that we have determined is valuable.

So, my real question is, we talked about periodic monitoring, what does that mean? How do you monitor? How often do you monitor, because that becomes crucial in identifying whether or not there are any declines over time.

CO-CHAIR KNOWLTON: Peter?

MEMBER SCHMIDT: So, I agree with Michael that this has been topped out for five years, and I think that the issue of -- I wish that people paid so much attention to what was endorsed by this Committee that we could reach 98

percent of patients with a measure. But, I think that probably this is being put broadly, and we have to beware of the late woebegone phenomenon of declaring that every hospital is above average.

CO-CHAIR KNOWLTON: Reuven?

MEMBER FERZIGER: So, I use this as guidance, additional guidance, on how to think about those all season patients of the reserve status, because it's hard to separate the observer phenomenon out of this.

Like in Michael's point, is that in this particular case, right, we already knew, but then it's not clear why we ever needed a measure in the first place, right?

But, assuming that there was a use you have a measure, right?

You are talking about the fourth leading cause of death. So, why wouldn't we err, if we think that there's a possibility, right, but there's an observer phenomenon of rising improvements, why wouldn't we want to err on the

side of over measuring, or continuing to measure, until we have some information, or, let's say, we have an eMeasure that replaces it, and takes away, you know, the user opportunity cost issue.

It seems to me like from a policy point of view, we would want to keep measuring this.

MS. JOHNSON: That, actually, that argument sounds familiar to me from the last time around, and it may sound familiar to some of you who were on the panel before. I don't know that with this particular measure, but I do know some measures the last time that we were getting towards the topped out level, the Neuro Committee the last time did say, well, if there still seems to be a little bit of space, and because stroke is such a high prevalent position, even that little bit translates to a lot of people. And, that was the argument for at least one or two of the measures the last time around to go forward with it.

MEMBER FERZIGER: It seems to me a

little bit different, though. It's not just that there might be even a very small gap, but that's a significant number of people, I'm saying that if we have any information about how much, you know, the sustainability, you know, of the measure, or, you know, or of the clinical phenomenon depends on the existence of measurement.

CO-CHAIR KNOWLTON: Jim.

MEMBER BURKE: This is sort of follow-up on the criteria as well. I mean, the first one also seems like it's one that's not super obvious, about, like David, how you make that judgment. So, what we know is the number got better. But, do we know sort of the things like why it got better, is it documentation commonly exclude quality phenomena, and that seems like it's hard to get at.

I can imagine theoretical ways that could be measured or gotten at, but I don't know that we have any measurements here and I don't know in the future if you make any reserve

judgements that might be helpful to know.

I mean, what proportion of this -- are we seeing changes in documentation prior to the example of the phenomena? Are we seeing something else correlated here? I mean, I think with this measure it's hard to know which of those is, I hope, with accurate importance, but I don't know.

CO-CHAIR KNOWLTON: Charlotte?

MEMBER JONES: When we ask people to be accountable, when NQF says this is something we feel that you should still be accountable for and you need to measure, something else is not being measured.

We, certainly, understand the logic that stroke is a huge population and a huge problem, but everyone lives in a world of limited resources. And so, even if we put something on reserve, or keep it active, we are taking resources from something else.

CO-CHAIR KNOWLTON: Okay. Ketan.

MEMBER BULSARA: Just to follow up on what David and Reuven have said. There's

probably historical -- there's probably historical data on what happens to these measures that are either retired or not continued to be endorsed by NQF.

And, I think having a sense for what happens to those measures in terms of compliance once that status has been established would be very useful in helping understand whether this is something -- whether it's a transient victory or whether it's going to be a permanent victory.

MEMBER HACKNEY: That would depend in part on the extent to which this measure is responsible for the improvement.

But, as I said, you know, get with the guidelines I don't think was just one measure, it was an overall effort to get people, actually, perhaps, in accordance with this. This was part of it, but I don't know of any way we could figure out how much of it. I guess if nothing else changes this goes into reserve status, compliance goes down, then we say, ah-ha, that was the lynch pin that made the whole thing work.

But, there's no way to know.

CO-CHAIR KNOWLTON: Karen.

MS. JOHNSON: I'm sorry, there's a little bit more of the policy that I thought I should read to you, and I think we touched on it before, but let me just read it out.

"Measures assigned to the inactive endorsement status will not be reviewed in the usual endorsement maintenance review cycle. During portfolio review, the Standing Committee will periodically review measures in the reserve status for any change in evidence, evidence of deterioration in performance or unintended consequences, or any other concerns related to the measure. The Standing Committee may remove a measure from inactive endorsement status if the measure no longer meets NQF endorsement criteria. A maintenance review may occur upon request from the Standing Committee or measure steward to return the measure back from reserve."

CO-CHAIR KNOWLTON: Go ahead, Michael.

MEMBER KAPLITT: So, with all due

respect to the NQF people here, I personally don't feel like we are getting adequate guidance from you guys, because I don't think that any of us are, actually, disagreeing on any of the things we are talking about. We all have our own interpretation of what process we are trying to follow here, and I think that's the problem, you know, you need greater clarity in my view.

I think if we are here to say is this important or is this not important, I don't think any of us disagree that this is important. If we are here to say that, you know, should we continue as a group to continue to review periodically on a set schedule to set criteria this measure, and that's the difference between what we are arguing, that's a very different thing. It doesn't mean it's not important.

So, for example, if we put this into reserve status, JCAHO is still going to use this, right? Yes?

Well, you know, this is I think the problem a lot of us are having, because we don't

know the consequence of what we are doing right now, I think. We are all interpreting it for ourselves. I mean, maybe I'm the only one who is not sure, but the sense I get is that we all have a different idea of what's the consequence of the vote we are about to take. We don't have any disagreement on the importance of doing this in clinical practice.

So, I personally think we need better guidance on what the consequence of this vote is, because otherwise we've got 23 different opinions as to what we are doing.

CO-CHAIR KNOWLTON: Jane?

MEMBER J. SULLIVAN: So, piggybacking on what you said, Karen, where if the measure goes into reserve status, and the Committee at some point decides that they want to look at ongoing data, who is responsible for generating that data, because at this point the developer provides the data to the Committee. But, if there is -- is there no longer a developer if the measure goes into reserve status, and we say we

want to look at data, where's that data coming from?

MS. JOHNSON: Let's be clear that if it is in reserve status it is still endorsed, so our developers have so agreed, unless they change their mind, that they will support the measure. So, we would expect that if we asked for it the developers would give us the data.

CO-CHAIR KNOWLTON: I'm going to suggest we take this up after lunch. I'm going to sit with it. I would like to have a few minutes with staff to talk about this a little bit, and see if we can at least get some clarity to what we are voting on.

I think, Michael's point was well taken. We need to be clear on what we are voting on. We might not agree with it. We might, whatever, but let's at least be clear on what it is.

So, let's take some time to clarify this. We are clarifying it by questions and not by answers, so let's get some -- let's have the

discussion with staff during the break.

Let's take a public comment now, NQF member and Public Comment now, have our lunch break, see if we can clarify the waters a little bit, and bring it up right after we return.

So, if the operator would open the lines for member or public comment, and I would, while she's doing that, make the comment that the comments by the public, and this will include public present in our room here, that they should comment on the issues that we've had before the Committee thus far, not anything that we haven't taken up yet.

Operator, are there any public comments?

OPERATOR: If you'd like to make a public comment at this time, please press star, then the No. 1 on your telephone keypad.

There are no public comments.

CO-CHAIR KNOWLTON: Thank you. Anybody in the room?

Seeing none, we are going to go to

lunch. We will be reconvening at 1:10, 1:15.

(Whereupon, the above-entitled matter was recessed at 12:44 p.m., to reconvene this same day at 1:15 p.m.)

CO-CHAIR KNOWLTON: Okay. I believe we're all back.

We're going to revisit the reserve status and David gave a very good summary. Gave absolute pristine clarity to this issue.

CO-CHAIR TIRSCHWELL: You're setting me up for failure there.

All right. So, we're voting about reserve status versus not. And to be clear, not reserve status means we would be not endorsing this measure. It would no longer be endorsed by NQF and as far as interpreting how to feel about those things, if you put something on reserve status you might want to do that because, yes, we're achieving great things but we're worried that if we completely unendorse it we might lose ground and lose quality of care. So, we want to keep an eye on it. So, we're going to put it on

reserve status versus if we fully don't endorse it then, in fact, we're so confident that this quality is maximal and will be maintained that we just really don't need to put this limited resource for quality measuring and improvement into this particular area at this time. So, that's the vote that we're making here.

CO-CHAIR KNOWLTON: Okay. So, we are considering measure 0435, discharged on antithrombotic therapy.

Comments on reserve status? Nobody has a comment so everybody is ready to vote? Okay. Christy, you're handling the vote?

MS. SKIPPER: Just give me one moment.

CO-CHAIR KNOWLTON: Did you vote for us, Christy?

CO-CHAIR TIRSCHWELL: So, the question is, does the Chair have any recommendation? Well, you know, that's a great question. I do know how -- I guess you're asking me how I'm going to vote and why. And, honestly, I am actually leaning towards more reserve status for

this because I do worry that if we completely take it off the radar, that we might fall back and I would like that additional information.

I think in the future on re-evaluation we could always remove it from endorsement later as well. Is that correct? So, that's sort of where I am.

MEMBER HACKNEY: The only thing I'd bring up is I think we've made the suggestion that someone needs to address this issue that the eMeasure can't resist because going forward even if this is on reserve status or even if it was still active nobody is going to do it anyway other than electronically. And if it's being done electronically then the overhead is pretty low. So, ultimately you'd like to have a measure that is usable, that's electronic, but right now we don't have a route to get there.

CO-CHAIR KNOWLTON: I think it's fair, David, just to highlight on that that we need to be thinking also through the implications of getting rid of an abstracted validation

alternative because I think as Lee was commenting during the break, how do you measure people who do not get care and how do you validate your data if you don't have access to the abstracted record to be able to do that?

So, it seems like the NQF part of our thoughts going back might be that NQF has to give it some thought to what the implications are because here's a measure that is clearly topped out. I mean, I don't know many measures that have got this type of compliance with it. But the reality is so it's almost as if not this one which one would you say we're done with this one?

On the other hand, what are the implications and that's the piece that I don't believe has been thoroughly vetted, thought through. But in this imperfect world we are called upon to vote.

So, anybody want to -- a burning need to make one more comment or can we vote on it?

Okay. Are we ready to vote? Christy? Not yet.

MS. OGUNGBEMI: Polling is now open. We're voting on whether or not to move 0435 to reserve status. One yes, two no.

Poling has closed. Eighty-seven percent, yes, thirteen percent, no. The committee has voted to move 0435 to reserve status.

CO-CHAIR TIRSCHWELL: Okay. So, we're finished with that. Right? We can move on to our after lunch agenda.

And so as we're moving into the next set of measures I would note that it's three more stroke measures from the Joint Commission and each one has a companion eMeasure and so some of these same issues which we've struggled with and mastered may come up again so we'll see how we go timing-wise and I guess maybe we could pass the baton to the Joint Commission and Dr. Lee Schwamm is here now as well representing the Joint Commission to let you guys -- why don't you just give us the overview of the first measure which are the two statin ones. We'll do it one at a

time so we don't have to remember too much.

MS. KOLBUSZ: Well, thank you for introducing Dr. Lee Schwamm. He's our physician consultant for this measure as such.

The next measure which is stroke 6, discharged on a statin medication, also last endorsed in 2012. This measure captures the proportion of ischemic stroke patients who are prescribed a statin medication at hospital discharge.

The denominator population for this measure has changed since the last endorsement period. It is now all ischemic stroke patients. It was updated to align with an update in guideline recommendations that occurred from ACC and AHA in November of 2013.

CO-CHAIR TIRSCHWELL: Great. So, then I think I'm on deck for talking and I just want to find my notes here. All right.

So, David, you're going to lead us through this. I'm going to do some of the talking. So, I'm going to start with evidence.

Right?

CO-CHAIR KNOWLTON: Right.

CO-CHAIR TIRSCHWELL: So, there hasn't been any new clinical trial evidence since this statin measure was first approved when we were last here. Some of the details of the numerator and the denominator have changed a little bit to come into alignment with this new guideline that was already referred to. But the fact that statins are effective and important at reducing outcomes after atherosclerotic-related ischemic stroke that's the evidence that this is based on and, you know, I don't think there's much debate about that. And so, you know, the preliminary ratings and I'm a little -- on my form there's two versions of it. Oh, it says moderate and then fast and I certainly think it's at least moderate, although maybe there's a reason why it can't be higher than moderate. Quality high, quality moderate. I'm not seeing necessarily a reason. I'm just thinking about the charts.

Like the other one, there was a reason

there wasn't something that prevented it from getting high. Is that in place here as well?

MS. JOHNSON: No, I think this one probably the staff landed on moderate probably because of the level B evidence. It's probably --

CO-CHAIR TIRSCHWELL: Okay.

MS. JOHNSON: But it's not like for reliability or something where we're saying testing is a certain level so moderate is as high as you can go, it's not like that.

CO-CHAIR TIRSCHWELL: And high or moderate I think there's little debate that there's clear evidence to support this. So, I -- that's all I've got.

CO-CHAIR KNOWLTON: Do we have any questions or comments?

Are you ready for a vote? Oh, so we don't need to vote because there's no new evidence. Okay.

CO-CHAIR TIRSCHWELL: I'm fine. I think we would have passed the vote anyway. So,

it's moot.

CO-CHAIR KNOWLTON: Rather moot.

CO-CHAIR TIRSCHWELL: All right. So, then we're down to our next criteria, the gap. And that same table that was in some of the earlier measures that caused consternation about whether the gap is topped -- whether there is a gap, whether we're already topped out on this measure. So, just looking at the data -- are you guys able to bring up this table? The one with the pink header line.

Looking from 2010 to 2014, the 50th percentile went up from 94 percent to 99 percent. The 10th percentile went from 71 percent to 90th percentile. So, you know, clearly this one is not quite as topped out if that's a reasonable way to describe it as the anti-platelet medication was. That being said, things are looking pretty good now. The 10th percentile only takes us down to 90 percent.

As far as disparities data go, it's the same thing, there wasn't -- I'm sorry? It's

on the other screen now. That makes it pretty much impossible to see. Sorry.

There was a generic reference to disparities but there actually were some more specific examples from the literature given. I'm guessing there's a possibility that we may see more data later but, again, the decision is whether the difference, you know, between the 10th and 90th percentile now is 90 percent versus 100 percent, whether even if there were gaps, whether that still represents a substantial enough gap to keep it fully endorsed.

And so I think we all understand these issues. We'll open it up to comments from the group and then we'll give the developer a chance absolutely before we vote. Is that a reasonable process?

CO-CHAIR KNOWLTON: Yes. Questions? Comments? Lee or Karen?

DR. SCHWAMM: Sure, I think it's important to note that the data you see on the performance gap is based on the prior denominator

exclusions where it applied only to patients who had an LDL greater than 100. The new denominator because of the subsequent evidence and the upgraded guidelines from the American Heart Association now include everybody with an LDL down to 70. So, there's a large number of patients who had moderate LDL elevations who are now eligible for the measure and we don't know the performance gap in those patients. Isn't that correct?

MS. KOLBUSZ: Correct.

CO-CHAIR TIRSCHWELL: So, we're presented with this measure where we've all agreed the evidence hasn't changed. The specifications of the measure have changed. We have no data about how the new specifications work and we're trying to decide whether there's still a gap. I don't know how we'd come to that conclusion.

MS. KOLBUSZ: I'd just like to clarify based on the opening of the project and the close of this project. No data had been received using

the new denominator population. So, there is a four month lag. The change in specifications went into effect October 1st of 2015. And those data had just been received at the Joint Commission. It would only be one quarter of data. So, it is impossible for us to provide you with data reflective of the change in the denominator population. There's just not enough time to do so. We need to reflect more.

DR. SCHWAMM: Okay. But it is a significant expansion of the nominator. So you'd have to hypothesize that the performance increased to the same proportion as the denominator expanded. That's the key point here.

Many more patients are eligible for this measure than were before.

CO-CHAIR TIRSCHWELL: I guess as it seems to be the case again and again we're stuck in a little bit of a policy issue about how do we evaluate this measure and the gap with the possibility that now that the specifications have been changed this older gap data which suggests

that it's pretty close to topped out if not already there may no longer be the case.

I mean, what are our options? Again, should we table this and wait for data? Should we judge it now but then potentially look at the data when you have a chance to gather some? I guess to do that like the earlier one, we would have to vote that it was insufficient which would give the Joint Commission the opportunity to come back with some data that would allow us to evaluate this particular question. Is that correct?

MS. ISIJOLA: So you're looking at the measure as specified today if in fact there is additional information, we can present that to you in working with the Joint Commission to review it during our post-comment call.

CO-CHAIR TIRSCHWELL: So, let me just backtrack a little. There must be a standard process for approving modifications to a measure. What is that process? Is it different than what we're doing right here?

MS. JOHNSON: I'm sorry, you caught me in mid-chew.

Yes, we actually do have a maintenance process. We call it the annual update. So, every year we ask our developers to update their measure. So, if they've made changes to their measure we ask them to tell us about those changes. Often those might be as was discussed earlier a new med is on the market and so they update med lists, those kinds of things.

We look at the updates that have been made and we decide at that point whether the change has been a big enough change that we actually have to go through what we call maintenance process. So, their change in their measure just happened to coincide with our project that's happening now. So, we're not going through an ad hoc which would be our way of looking at a change that's happened kind of outside our cycle. We're just looking at it right now.

So, did that answer your question,

David?

CO-CHAIR TIRSCHWELL: I guess I understand that but it seems like the timing is undermining our ability to do so in that the timing didn't allow you guys to have any data to present. We understand that. But yet we're trying to evaluate the gap. I guess I'm feeling like there is insufficient information for which me to evaluate whether there's a gap or not.

CO-CHAIR KNOWLTON: Valerie?

MEMBER COTTER: Can I just make a comment? If the guideline came out in 2010 this is six years later.

MS. KOLBUSZ: The new guideline was from ACC&H in November of 2013.

MS. ISIJOLA: 2013.

MS. KOLBUSZ: And that's when it was first released. And then the following, I believe, May or June 2014, the secondary prevention guidelines from the American Heart were updated that reflect the same. But during that first year period there was discussion with

our technical advisory panel members. They wanted to see, you know, how this was assimilated, I guess, and we didn't make the change right away. And then it does take some time once we decide to make a change because the specifications manual is actually delayed. So, this change was really decided about a year ago but didn't go into effect until October 1st of 2015.

CO-CHAIR KNOWLTON: David?

DR. SCHWAMM: The only data I can give you, I mean, I can't give you exact data because I can't actually run this measure. But get with the -- I just did it right now. And get with the guidelines we have two parallel measures. We have a measure that you saw four years ago which was if you have an LDL of 100 or greater than the LDL for the measure, what's the frequency of adherence and it's basically 94 percent.

If you run the parallel measure which is, did you have a stroke due to atherosclerosis and get high intensity statin therapy so it's not

all statins. It's high intensity statins therapy. That's only about 53 percent.

So, in the group with the larger denominator but requiring high intensive statins, not any statin, the adherence is only about 54 percent.

CO-CHAIR TIRSCHWELL: So, that more accurately reflects the new specifications, is that what you're suggesting?

DR. SCHWAMM: That reflects the new denominator. The new numerator is still statin therapy, not high intensity statin. So to be in the numerator of this other measure but it's not perfect. I'm just giving you the best that I have. It's only about 54 percent.

CO-CHAIR TIRSCHWELL: Okay. Peter?

MEMBER SCHMIDT: So, I'm a little bit confused because there's no new evidence and the numerator and denominator are definitions of change. So, we didn't review evidence but there's been a change and it's not within the measure that was approved previously.

CO-CHAIR TIRSCHWELL: Yes, so I struggled with this and about how much detail to go into when discussing all of this. The previous measure of specifications was sort of, you know, straight out of the one clinical trial that looked at this and, you know, I think, reasonably guidelines have assimilated a greater body of information to expand the denominator a little bit more so it really isn't -- I still think there isn't any new evidence but the guidelines have changed a little to reflect a broader evaluation of who this type of treatment is appropriate for. But I feel that tension between those two things. No new evidence so why are we changing the specification?

MEMBER SCHMIDT: But wouldn't that constitute like class D evidence, expert opinion, if it's a new guideline?

MS. KOLBUSZ: The guideline recommendations still, class 1, level of evidence A, which is what it was before. Where the change is that there wasn't a change in the level of

evidence. They opened up the guideline recommendation to include all ischemic stroke patients because they identified in the new guidelines four statin benefit groups. And those with clinical ASTBD are put in the first group recommending high statin therapy and your stroke and TIA patients fall into that group.

DR. SCHWAMM: Yes, I mean I think this is a -- this was a decision by the experts that stroke should be thought of as an atherosclerotic disease and not a spec separate carve out. Prior to that there had been no data to the benefit of statins and stroke. The SPARCL trial definitively showed that high intensity statin therapy dramatically reduced the risk of recurring stroke and of cardiac disease, coronary events and coronary revasculation.

So, what the new guideline did is it basically said, look, we know from the data that everybody benefits who has atherosclerosis from statin therapy and so hundred is an arbitrary number. We should get rid of that exclusion

because if you have an LDL of 99 you know you still benefit by reducing the LDL. There's no drop off on that curve. There's no step function.

So, there isn't new evidence. You're absolutely right, but the decision was to broaden the impact of statin use to anyone with evidence of atherosclerosis and so stroke due to or in the presence of athero is really reconceptualizing that patient as a patient with atherosclerotic disease and, therefore, should be on statin therapy.

CO-CHAIR TIRSCHWELL: And I'll just comment that in the 2014 Secondary Stroke Prevention Guidelines there are three parts to the lipid thing. I think only one of them is 1A evidence. There's a 1B and I'm not going to remember the rest accurately so I won't comment.

But, anyway, suffice to say the evidence is still pretty strong but the specifications have changed. And I think the changing specifications are probably more

important about what it says and what we don't know about the gap now then it is about whether there's evidence to still give these patients statins.

MS. JOHNSON: So, this is kind of an interesting conundrum that we're in and we actually have a potential way out of this. We might not like it. But what we could do, is you have the option as the steering committee to defer your consideration of this measure. And by defer, we would ask you to kind of a joint agreement with Joint Commission. When do they think they could actually have some data because what's happened is it's a change in specification that's actually made the data available not quite what you need to be able to look at.

So, that would be an option. So, to make it clear you could ask to defer complete discussion of this measure and Joint Commission you can decide, is it's six months or a year or whatever is a reasonable amount of time for them to be able to have data that reflect the new

specifications.

MEMBER RAE-GRANT: A couple points.

First of all, I don't think we need to tie ourselves in knots around this. I think it is within our purview, I'm putting this out there, that if we perceive from the evidence that's not necessarily right here that there may be a gap still within the -- we probably don't need to vote against it if we think there's more to learn. And I don't know if we need to slavish about what's written down in the preliminary document in terms of our deliberations. So, that's one thing.

The second is I would think about -- there's different process measures and some of them are non-treatment process measures. I'm thinking of NIH stroke scale measures. I would think we should have a higher level of stringency to retain treatment-based process measures such as institutional hydro statin or whatever statin. Then we do for measures which are just measuring times to, you know, CT scan or something. And,

therefore, if we have that then we would want to retain a measure like this longer to insure that there's high compliance before we put it on reserve status. Does that make sense? So, we're only one year into or a couple of years into a final -- we're saying it's in compliance fully. Maybe we need to give more time to medication treatment or some other treatment process measure and that way both of those would help us not worry about this. Just vote it through and see what happens the next time we review it.

MEMBER FERZIGER: So, I agree with that completely. I have a question though about sort of where the state of either the treatment development or population definition is in this field, right? So, now you've made a modification to the scale. Right. What's the likelihood in two years, right, that either the population, you know, will be what we now have specified or that it would assimilate, right, be expanded because of new knowledge or that new treatment development actually would change, you know, the

focus from statins.

CO-CHAIR TIRSCHWELL: I mean, things can change. And there are processes for that.

MEMBER FERZIGER: So, I guess my thought would be, right. Everything is always changing and it seems that as long -- until things are stable and likely to stay that way for a long time there's no such thing as topped out because, right, because you actually have more things to measure that you haven't measured before and that are now important.

CO-CHAIR TIRSCHWELL: Just by the virtue of the fact that there is a change perhaps we should leave it in because --

MEMBER FERZIGER: Yes. If it's a substantial change, right, and that needs to be covered as well, the measurement.

CO-CHAIR KNOWLTON: Other comments?

MEMBER BULSARA: You know, I think I'm with David in the sense that we don't want to under serve this measure in the sense that we don't to vote on it one way when Lee is telling

us there's a lot more data so I mean just listening to this, I think the option to deferring sounds very appealing. But I don't know what the implications of that are but it does sound very appealing.

CO-CHAIR KNOWLTON: David.

MEMBER ANDREWS: As sort of an outsider to all this it strikes me that a lot of things change fairly quickly in modern medicine. And one of the things that NQF probably needs to grapple with is how do you have ongoing processes to deal with things that change much more rapidly than standing committees meet in order to arrive at a process that's responsive to the reality out there in the field?

CO-CHAIR KNOWLTON: Other considerations? So, it seems that we've got before us either if I understand correctly, correct me if I'm wrong. We have either a vote on gap or we have a vote on deferring. Is that fair?

CO-CHAIR TIRSCHWELL: Well, I think if

we vote on gap and you want to defer you should probably vote insufficient. Is that right or is that a different thing altogether?

MS. JOHNSON: Maybe let's get a gestalt first. Let's just see if anybody has, you know, we've heard one person interested in a deferral and some other folks probably not so interested in deferral. Can we just do a show of hands and see?

CO-CHAIR TIRSCHWELL: Well, I think -- I think there are a number of options. One is, you know, as Alex was saying or actually, I'm sorry, somebody, you know, things have changed and common sense says to me that this is going to open gaps, not close them. And so I'm willing to just say that their gap is moderate and move ahead with continued endorsement. That would be one options.

Another option would be that I've got data in front of me and, you know, I don't -- I don't have any data about this new thing. I don't think it's going to be so low. So, I just

want to say that there is no gap left and it's low.

Or that I just don't know and I want to see those data from the first few months and that would sort of be the insufficient argument.

Or do you want to add a fourth one to that?

MS. JOHNSON: I don't want to add a fourth one but I would be curious from TJC when do you think he would have some data to be able to show because our post comment call is going to be, what? Two months from now? So, I guess the question is, will there be anything two months from now for you to look at?

MS. WATT: Well, our question right back at you is, how much data do you think you need? We know that we have one quarter of data now that just came in or is just coming in on April 29th and it's going to be another four months until we get the data for first quarter of 2016.

If one quarter of data would be

sufficient for you we should be able to share that sometime early May.

DR. SCHWAMM: Just -- I don't know if it will come in in time, but I just actually sent an email out to someone at AHA who is running the measure with the guideline hospitals. And for what we have for 2016 so it may be that in 10 or 15 minutes so I can give you a rough estimate of performance in the first three months of the year.

CO-CHAIR TIRSCHWELL: I mean, get with the guidelines, collects the newly specified measure as a conduit to data submission to the Joint Commission. Did I use all those word correctly? And get with the guidelines. You know, the numbers are staggering so three months worth of data is going to be thousands and thousands of patients undoubtedly, right? So, that would seem adequate to me. I would know a lot more then than I do now. It would potentially be a bias towards better performing hospitals so we might imagine them nationally.

CO-CHAIR KNOWLTON: Which may show that gap much narrower, you know, because get with the guidelines hospital are the --

CO-CHAIR TIRSCHWELL: Well, there's a lot of them out there. They're not all perfect.

CO-CHAIR KNOWLTON: So, what was your recommendation, Karen, that we do what?

MS. JOHNSON: Well, why don't we -- it sounds like we might have information in 15 or 20 minutes. We might have it in May that would be a quarter and it sounds like that would be a lot of data, right? Why don't we -- I don't know, Elisa. Should we stop discussion of this one. Go to the next one and then come back? Or do we wait? What do you think?

MS. MUNTHALI: Let's stop and go to the next one.

MS. JOHNSON: Okay. So, we'll see what -- give us guidelines. Your analyst is back there --

CO-CHAIR KNOWLTON: Charlotte, let's see what Charlotte --

(Off mic comments.)

CO-CHAIR KNOWLTON: I think it's a little different. They have some data. The difference is there was an actual change in the standard and they've got data that supports that.

When you talked about disparities they didn't have that data yet. And they're saying, this is what we think will happen. That's different than we have data that you can look at, right?

MEMBER JONES: But isn't the process the same? Is it one time we're saying we don't trust the data that's being reported to us in this meeting and now we're saying we do trust the data that's being reported to us in this meeting.

CO-CHAIR TIRSCHWELL: Instead of saying that we don't trust it --

MEMBER JONES: And I'm not.

CO-CHAIR TIRSCHWELL: We just don't have the written data to fully review and take time to assess. But I think your point is well taken. That's even when Dr. Schwamm who adds a

new lightning fast data analytic element to the Joint Commission but still won't have it in front of us to evaluate completely and ponder without feeling a time pressure.

MEMBER JONES: I think in terms of transparency we weren't willing to accept data provided to us in a previous discussion. To do so for this raises a transparency issue. And I just think we need to be aware of that.

MEMBER FERZIGER: So, I just want to emphasize what you just said, right? That maybe we can project what data we will have, let's say it's beautiful data, you know, and it's very good for where it comes from. But still, you know, it leaves the gap, right, of, you know, the range of places it could come from. There's likely to be some difference. So, that means that, you know, even if we imagine the best case scenario, some of us will still think in three months that there's a gap. So, I would be prepared to vote on that basis that even if I have everything I got or I wanted I'd still think there was a gap for

the reason you just identified.

MEMBER BULSARA: Just a follow-up on what Charlotte was saying. I think in all fairness to us, I think we do have to have an opportunity to actually look at the data and in fairness I think the Joint Commission, I mean, they should -- if us deferring this doesn't result in loss of endorsement over the next couple of months, I think they should have a fair opportunity to present the best data. So, I think it works both ways and I mean if there's no penalty to deferring, I don't see why we should rush through it.

CO-CHAIR KNOWLTON: I agree with that.

CO-CHAIR TIRSCHWELL: Karen, are we back to taking your straw poll on deferring.

MS. JOHNSON: I think we might be. There would not be -- they would not lose endorsement. It would just be deferred.

MS. MUNTHALI: And I just wanted to clear up one thing about the difference between this measure and the other. On the clinical

practice side, guidelines changed for this measure which was very much out of the control of the Joint Commission in terms of the timing. We were trying to catch up. The deferral option that the committee has -- one of the reasons you can defer is the timing of guidelines. So, that is very much within our process. I did want to clarify that and that's why we'd rather you defer if at all possible. The measure isn't going to lose endorsement but it gives the Joint Commission enough time to catch up with practice guidelines.

CO-CHAIR KNOWLTON: Other comments?

Shall we -- what do we do a straw poll, David, you think for the --

CO-CHAIR TIRSCHWELL: Show of hands?

CO-CHAIR KNOWLTON: Show of hands? We don't have a way to electronically vote for that?

So, how many would be in favor or deferral of this measure now? Does anybody oppose it? Okay. So, because this is deferred we move on, is that correct?

CO-CHAIR TIRSCHWELL: And we'll also be deferring the eMeasure then I imagine as well.

CO-CHAIR KNOWLTON: Right. We'd have to.

CO-CHAIR TIRSCHWELL: Yes.

MS. WATT: Excuse me, this is Ann. I just need a little clarification.

So, is the one quarter of data -- will that be sufficient for you all to be able to make this determination and were we going to discuss the rest of the criteria as well as we did earlier when we -- I'm just --

CO-CHAIR TIRSCHWELL: When we re-review it after we have the one quarter of data we'll go through the rest of the criteria as well.

MS. WATT: Okay. And so one quarter of data will be sufficient?

CO-CHAIR TIRSCHWELL: You know, I guess there's a possibility that we would say it's not. But given the numbers, I mean, it's hard for me to believe that you wouldn't be able to demonstrate a gap if there is one there.

Okay. Moving along. We're going to go to the thrombolytic therapy measures 0437 and it's e-companion 2834.

Would you like to introduce the measure?

MS. KOLBUSZ: This is our stroke four measure, which is our thrombolytic therapy measure, the measure that captures the proportion of acute ischemic stroke patients who arrive at this hospital within two hours of time last known well for who IV-tPA was initiated at this hospital within three hours of time last known well.

The rationale is supported by an in-study. The evidence has been well established that the early administration of thrombolytic therapy to eligible ischemic stroke patients within that three hour time frame does actually improve neurological outcomes for patients. Time is considered brain so earlier administration rather than later is preferred, although more recent guidelines from the European Cooperative

Acute Stroke Study III did show that it can be administered safely and effectively up to four and a half hours which we acknowledge. But we still do maintain that the high bar is the three hour administration time frame.

CO-CHAIR TIRSCHWELL: Great and our discussants are Rubin and Mike. Who is going to take the helm?

MEMBER KAPLITT: Yes, I will. So, hopefully this will be less controversial than the last couple but who knows.

So, the evidence for this as you heard has not changed very much since the last endorsement. The numerator is basically, you know, all -- is patients who arrive within three hours or last -- who receive IV-tPA within three hours of last known well and the denominator is basically all patients who are eligible with some exclusions that we'll talk about under validity.

The evidence in favor of that three hour window hasn't really changed much as you just heard. I mean, there's a wealth of evidence

that supports that the three hour window makes sense from, you know, randomized studies, NIH studies that the FDA approval of tPA, etcetera, etcetera. So, I don't think there's new evidence that contravenes any of that to make it controversial.

The only issue is this European study that shows that you can in some patients extend the window to four and a half hours but I don't think that that materially changes the evidence in favor of this guideline. It really relates more to the validity question that I'll come to later. Unless, you know, people feel strongly that all patients should be treated within that four and half hour window and I don't think the one European study will support that.

CO-CHAIR TIRSCHWELL: Well, again, I think we're judging the measure that's before us.

MEMBER KAPLITT: No, what I'm saying, they'll change their measure to include that. So, I'm saying on the evidence -- so they've created a new exclusion criteria based on that

three to four and a half hour window of documenting patients who have -- who receive IV-tPA outside the three-hour window. That's new since the last measure. It's not part of the inclusion, it's a new, whatever data element.

CO-CHAIR TIRSCHWELL: Yes. Clarify that a little bit.

MS. KOLBUSZ: I just want to clarify. That is not new. The exclusion to exclude patients who receive tPA administration up to four and a half hours was possible with the last re-endorsement in 2012. The way we went about it though was different at that time. We relied on text documentation instructing the hospitals through abstraction guidelines that if they had a reason for not administering IV-tPA within three hours that they were allowed to select no for IV-tPA administration so that they would be able to reach the reason, data element in the algorithm.

In working with CMS and the data warehouses, that logic might work for humans. It's backwards computer logic and we needed to

correct the algorithm logic so that we would accurately categorize and capture those patients. Therefore, in working with the CMS technical work group we did the revision and added a very clear data element for exclusion of patients and we call it reason for extending the initiation of IV thrombolytic to exclude those patients who receive IV thrombolytic therapy within three to four and a half hour window when they have a valid reason for exclusion. But -- yes, they still have to come in with two hours. So, we basically recut our timing too in the algorithm making a repeat decision since the reason is often the same that we repeated it and it's all really about computer logic and computer programming, the reason for that change. It isn't a clinical change per se.

MEMBER KAPLITT: Yes, okay, fine. I mean there's no point in wasting a lot of time. It's not a clinical change to the extent that you always had the ability to exclude patients for some medical reason. You've just created a new

data element that specifies this particular reason, right? Isn't that true?

MS. KOLBUSZ: The new data element captures that time frame three to four a half hours.

MEMBER KAPLITT: Right, but people can still exclude for other medical reasons too. Somebody gets a thrombectomy, for example, within two hours, right, they're not a candidate for IV-tPA. They can be legitimately excluded from this because that's a medical reason, right? But that's not got its own data element. That's a catchall of medical reasons. I mean, I think we're arguing a semantic point here. I'm just saying that I just want to make sure I'm being accurate on what we have here, right? You have a specific data element for extended IV-tPA as an exclusion.

MS. KOLBUSZ: Correct. Before there one data element reason for not initiating IV-thrombolytic.

MEMBER KAPLITT: Right.

MS. KOLBUSZ: Medical patient reasons would be excluded by that data element. Now we repeat it to try to accurately capture that time frame of three to four and a half hours.

MEMBER KAPLITT: Okay. So, my only point is that as far as evidence is concerned the only evidence issue that's raised by the European study is whether or not one believes that the three hour time window is no longer offered and should it be a four and a half hour time window because that is the basis for this study from an evidentiary standpoint. Forget about the validity which I think it relates to validity later on.

And my only point is that I don't think that that one study, you know, is sufficient to fundamentally change the evidence in support of this. I believe the evidence in support of this measure is still valid for three hours. That European study notwithstanding. I don't think that one European study that is cited here more from the exclusion standpoint rather

than, you know, as an evidence issue. I don't think that that changes the evidence materially. That's my only point.

CO-CHAIR TIRSCHWELL: Okay. Any other comments or discussion of evidence? And if not I guess if the evidence hasn't changed we don't have to vote on it again and we can just move to gap.

MEMBER KAPLITT: All right. So, the gap I don't know, can you put it up here? I mean, unlike the last measures I don't think that this is really topped out by at least the standards that we've been using. You know, you can see here that we're not anywhere close to the sort of 99 percent ranges and I think that is showing clear improvement which is good but I think there is still an obvious performance gap that you can see in the numbers here.

I think that from the standpoint of disparities as with all the other measures, there's no real disparity data provided. There's a lot of citations of disparities in stroke

outcomes among different, you know, ethnic groups and et cetera. You know, discussion of, you know, utilization and things like that. There were a few minor studies provided to suggest that the disparities were actually kind of narrowing a bit but no real data from this measure.

The only point that I would make is I think that, you know, the gap is still real and, you know, and needs to narrow. The only thing I would say to the developer is that given that we're not getting a whole lot of disparities data on many of these measures as this trend continues if we very soon in the next year or two start reaching into the 90 something percentile I'm just suggesting that we're going to wind up in a couple of years in the exact same situation we're winding up in the other measures and maybe we can do some preventative data collection to make sure that two years from now when this thing is reaching 95 percent, let's say, and we're arguing that there may not be as much of a gap anymore that we can actually see disparities data to see

whether that might represent an ongoing gap. We don't have it. I don't think we need it today because there's still an overall gap but I'm just throwing that out there.

CO-CHAIR TIRSCHWELL: And since we have Dr. Schwamm in the room and get with the guidelines is a big part of the data collection process, I know race and ethnicity and gender and age are all part of the massive database that is get with the guidelines and it seems to me it wouldn't be too terribly difficult to be able to break this down and look for disparities you have those data too? Okay.

MEMBER KAPLITT: But otherwise, I mean, so I think there is a gap. I don't know if anybody has anything to add.

CO-CHAIR TIRSCHWELL: Any other discussion around gap? Rubin? Yes, your card's up. No problem.

CO-CHAIR TIRSCHWELL: Okay. Let's move to --

MEMBER HUFF: Just a comment. As we

move toward licensing more individuals an unintended consequences are many individuals with stroke mimics are receiving thrombolysis and in most centers this ranges 20 to 30 percent is the information that I have and everybody agrees that thrombolytics clearly effective in a correctly selected patient. I think it would be interesting and probably you have the ability to gather data on thrombolytics administered to patients not with ischemic stroke but with stroke mimics.

CO-CHAIR TIRSCHWELL: That would be interesting data. I think if their hospital discharge diagnosis is not ischemic stroke then they're not going to be in any of these databases. So, I mean, those data are out there. I agree with your point that it's important that the high performing places have a 50 percent over-treatment rate versus, you know, 10 percent somewhere else. Then you'd maybe start worrying about it.

Any other comments before we go ahead

and vote? We bring up the vote. The screens are blank up there. Oh, here we go.

MS. OGUNGBEMI: One moment. We are now voting on evidence for measure 0437. Oh, pardon me.

We are now voting on performance gap for measure 0437. The options are high, moderate, low and insufficient. Voting is open.

The voting is closed. Results are 26 percent high, 70 percent moderate, 4 percent low and 0 percent insufficient. Measure 0437 passes on performance gap.

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

MEMBER KAPLITT: So, can we move on to reliability?

CO-CHAIR KNOWLTON: Reliability. Yes.

MEMBER KAPLITT: So, reliability just to the point that someone made earlier, I will say up front that only -- testing was only done at the data element level not at the measure level. So, theoretically, we should only be

voting moderate as the highest score. But I leave everybody to their consciences so.

So, at the data element level the vast majority of the data elements showed a good inter-related reliability above, you know, 90 percent for most of them. A few elements were above 80 percent but mostly in the high 80s. The only one that was a little low was time to last known well which was 80 percent in related reliability.

And then the only one element that was a little bit lower was the reason for not initiating IV-thrombolytic therapy which was 77 percent. So, I think it's reasonably reliable. I'm assuming but obviously we don't data yet but this new element might actually help with that last one that was a little low because it clarifies for people. It's a separate element. And one of the reasons which is giving them IV-tPA late but I think that these numbers still show reasonable reliability from my standpoint.

CO-CHAIR KNOWLTON: Michael, if you,

unless somebody objects, if you don't feel it's new information we don't need to vote.

MEMBER KAPLITT: Yes, that's fine with me.

CO-CHAIR KNOWLTON: Okay.

MEMBER KAPLITT: I think the validity thing I do have an issue to clarify.

CO-CHAIR KNOWLTON: Let's stick with -- then we'll move past reliability and validity. Go ahead.

MEMBER KAPLITT: Okay. So, I need to clarify something before I address validity with the developer.

So, there are two sets of data that I'm seeing in the measure. So, one is a set of data under -- sorry. There's one set of data that shows the exclusion percentages and then there's a new set of data since 2012 and they show very different numbers. Can you guys clarify what that is?

MEMBER KAPLITT: I can tell you the developer if you can clarify for me.

So, on page, for example, on page 33 of your measure under results for the potential threats to validity there's statistical analyses for exclusions and there's various percentages that are provided. But then several pages down where you're saying new since 2012 endorsement data for empirical testing there's a different set of numbers.

So, are the numbers on page 33 the old numbers from the prior measure and then what's on page 35 -- I'm sorry, not page 35, page 35 through 39, is that new data because those numbers are completely different and I'm trying to understand what they are.

MS. KOLBUSZ: I just have a comment to make. What you are using right now is the preliminary analysis that was prepared by NQF staff.

MEMBER KAPLITT: No, I'm looking at the actual measure. I'm not looking at the summary page. Well, I think I am. Maybe I'm not. I'm looking at -- yes, so 2B 3.3. There's one set of

data. And then there's a thing that says new since 2012 and which is a new -- essentially a new to be 3.3 and I don't know what -- I'm not understanding these numbers.

MS. KOLBUSZ: That is the measure worksheet. I'd like to just look at the testing form that we submitted to NQF --

MEMBER KAPLITT: Okay.

MS. KOLBUSZ: -- before I respond, please.

MEMBER KAPLITT: The problem is -- the reason that I'm trying to clarify this is because I don't understand what these numbers are because they're very different and they're two different sets of numbers for the same thing. So, I'm not sure where they're coming from.

I mean as an example, just to be clear of what we're talking about. So, it should be 3.3 and I guess it's the worksheet. I didn't realize that so I apologize because I thought the worksheet was this sort of short summary thing. But under 2B 3.3 where it says that out of 39,812

patients it says that the patients with documented reason for not initiating IV thrombolytic was .95 percent. Then under the thing that you were just talking about in red that you guys are showing here, it's showing for 206,000 patient records -- that's why I'm assuming that this is the difference between what was before and what was now but I don't know.

MS. KOLBUSZ: No, I think that now that I'm looking at the testing form that was submitted with the measure submission form from the Joint Commission 2B 3.2 asks, what were the statistical results from testing exclusions? And we did state that there were 2,206,379 admissions included in the initial cohort.

From among the 2,206,379 admissions and 13,018 hospitals, the descriptive statistics are given below and we provided exclusion data for the various data elements that result in exclusion. Clinical trials, elective carotid, time less known well was none. Time less known well to arrive in the ED greater than two hours,

none. And then also patients with the documents reason for not initiating IV thrombolytic and patients with a documented reason for extending the initiation of IV thrombolytic. That's one set of data.

On this form 2B 3.3 question. What is your interpretation of the results in terms of demonstrating that exclusions are needed to prevent unfair distortion of performance results? We stated the median frequency of exclusions range from low to moderate. The distribution exclusions across hospitals is generally narrow, indicating that the occurrence is random and likely would not bias performance results, although the percentage of patients excluded may differ depending on whether the hospital was a stroke center.

For criterion validity it's believed that all the exclusion should be retained for the following reasons. And then we provide rationale. There's no further data in the submission we provided.

MEMBER KAPLITT: So, maybe NQF can clarify for me what this other set of numbers is. I don't know who did this because the reason it matters is because if this other set of numbers was from the original thing from several years ago, it shows a massive change over time. That's why I need clarity where this other set of numbers come from.

So, if you add this so-called worksheet on page 33, 2B 3.3, it says an end of 39,812 patients, not two million patients, where does that come from?

CO-CHAIR TIRSCHWELL: Is that from before the work group?

MEMBER KAPLITT: I didn't think so but I can look it up online. I don't know.

Which I would like somebody to confirm. If that's correct which I would like somebody to confirm, but if that's correct then what it shows is a big shift over time in the percentage of exclusions as, you know, even though -- even though we've seen, you know, the

sort of performance gap reviews we've seen a much bigger -- a big change in the percentage of exclusion. For example, if this was the original data it showed .95 percent of patients excluded for documenting a reason for not initiating IV thrombolytic therapy and now it's like 20 something percent.

MEMBER SCHMIDT: The percentage in that section appears to be calculated incorrectly.

CO-CHAIR TIRSCHWELL: They used the wrong numbers. Is that right?

Are there two million -- there aren't, you know, there were not two million, I mean. There's a lot of records in get with the guidelines now what, two million something total or three million, four? Okay. But half of the patients in there weren't ischemic strokes that came in with two hours. So, the two million number should never have been -- that's the entire database probably at that point, right?

MEMBER KAPLITT: So, maybe not because

the carotid number says 11 percent. It says 259,000 patients which is about 11 percent of the two million patients.

CO-CHAIR TIRSCHWELL: Right. It doesn't make sense to me.

DR. SCHWAMM: I think the issue is do you select as your denominator all ischemic stroke patients or do you select as your denominator all ischemic stroke patients arriving within two hours? Right, because the reason for non-treatment varies depending on whether or not you're excluded based on the --

CO-CHAIR TIRSCHWELL: Well, what's the measure? I mean, you should select what the measure is, right?

MEMBER KAPLITT: It's your denominator.

CO-CHAIR TIRSCHWELL: The denominator is within two hours, right?

MEMBER KAPLITT: If the numbers would start making sense.

So, this is my overall concern with validity. I have a series of numbers that I

spend time trying to understand and I just decided that I had to come here to get clarification because I couldn't understand these numbers.

MS. KOLBUSZ: One thing that I could say in regards to looking at the previous data from 2012 now that that's clear on this form to me what you were referring to. If you would look back on the total number of hospitals that collected this measure, when we came for endorsement in 2012, CMS was not collecting these data. The only hospitals collecting these data were Joint Commission certified primary stroke centers. So, it's a much smaller number of hospitals and also probably your more high achieving early adoptive type hospital. The number of hospitals has really increased as you saw with the performance gap numbers when you look especially at years 2013 and 2014. There's a significant number of hospitals. So, it jumps a lot.

MEMBER KAPLITT: So, they do but if

this 24 percent is correct and I agree that it doesn't fit. If you take 73,000 over two million that doesn't fit. So, either it's a simple math error and I'm worried about something that's not real or your denominator is different here. But this number is wrong. But if it were 24 percent exclusion whereas in that original group it was only one percent exclusion then it does represent as you broadened out, well, the gap has reduced. It's reduced in part because had a much greater level of exclusion and then the question is, does that raise the validity question.

But we have a more fundamental issue here that has to be clarified which is are these numbers accurate because, you know, whoever spoke earlier, I apologize I didn't see, is right. They don't make sense.

CO-CHAIR TIRSCHWELL: Okay. Steve then Peter. Peter?

MEMBER SCHMIDT: So, your last comment was a little bit worrisome that the exclusions -- you seem to -- I interpret it as the exclusions

are used more extensively by the lower achieving hospitals and that sounds like a threat to validity to me.

CO-CHAIR TIRSCHWELL: So, I think there is some simple math problems here. There were some simple math problems with some of the other measures where 52 percent were being discharged home for hospice which never made any sense. And you've sent in some corrected data for that. I mean, the reality is, nobody is admitted for elective carotid intervention within two hours of their ischemic stroke. So that -- I mean, that doesn't -- it shouldn't even be part of this.

MEMBER SCHMIDT: I think this -- part of the challenge here and, again, I'm not familiar with the exact numbers that were submitted on this particular line. But the way the logic of the way the measure is constructed and which exclusions are applied in what order can have a big impact. So, admitted for elective carotid endarterectomy is an exclusion for all of the ischemic stroke measures because 433.10 in

ICD-9 is carotid stenosis or occlusion without visualized infarction. And so every carotid endarterectomy that gets admitted gets that ICD-9 code assigned to it. The problem is if you have a stroke due to your carotid artery but you have a pacemaker so you can't have an MRI and the CT doesn't show an infarct, those patients also appropriately are coded 433.10. So, the Joint Commission and CMS apply the carotid elective admission for carotid endarterectomy across all these measures before any other attributes of the measure exclusions are applied. So, that's why you see carotid endarterectomy showing up here. That's just giving you a sense of how many carotid endarterectomies were performed in patients who were discharged with that diagnosis code of 433.10.

So, I think part of the challenge here is the sequence in which the denominator exclusions are applied and that makes it hard unless you're looking at the actual raw data from the outputs to figure out in what group does the

proportion of exclusions seem to make sense or does it seem a little bit strange and it has to do with who has been filtered by the previous exclusion, not all applied at the same moment. I don't know if that is helpful.

CO-CHAIR TIRSCHWELL: That make sense to me even if some of the more common sense things that we're seeing don't make a lot of sense. And I guess it would be -- would have been nice for those data to have been clarified with the submission.

So, my -- sorry, Jim.

MEMBER BURKE: So, that reason is documented. There's no documentation specified in this exclusion criteria so that there is a reason documented but there's no judgment about whether or not that was a good reason, is that right? Do you understand that? Because that to me seems like, if I'm understanding that right, that seems like a -- or maybe I'm not understanding it right.

MS. KOLBUSZ: Could you rephrase that

because I'm trying to read and follow and --

MEMBER BURKE: No, understood. So, the question is the exclusion is that there was a reason documented or contraindication. But there's no specific criteria for evaluating whether or not the reason was a good reason or a valid reason, is that right?

MS. KOLBUSZ: That's correct. We do not pass judgment on what's documented. The documentation is taken at face value. We're looking for linkage with thrombolytic therapy so that we know that there was consideration of thrombolytic therapy but whatever was documented as the reason we would not be casting judgment with the exception of a few stand alone reasons that are in that data element which have been identified through working with our technical advisory panel. For example, if there was a documented miss of zero on arrival to the ED we would consider that a valid stand-alone reason and we wouldn't look for further documentation or linkage.

MEMBER BURKE: Okay. And that seems like it really amplifies Mike's concern if indeed these are real hospital level variations the hospitals were at 10 percentile is writing down a reason of 6 percent and the 90th percentile is writing down 54 percent whether or not we're judging how good you are at delivering tPA or how good you are at coming up with a documented exclusion, we're not going to hold those to a list. It seems like it's genuine validity concern. I'm not sure how much it changes things.

CO-CHAIR TIRSCHWELL: Okay. Thank you. And I guess just an observation. What's missing from this list of exclusions or maybe it's an inclusion but the patients that presented after 120 minutes, I mean, that seems like that would take a huge number out and that number is not listed here. I mean if we're taking out the ones that were electively admitted for carotid, sure. Those should, you know, they don't count but also the ones that presented later.

MS. KOLBUSZ: I'm sure it is a large number because when you look at the denominator population compared to other stroke measures, for example, stroke 2, you'll see that the denominator population for stroke 4 is always much smaller. It's only about a quarter percent. But that is based on a calculation. It's not based on a data element where we're reporting reliability of the validity for data elements, the calculation and the algorithm based on arrival date and time and date and time less known well.

CO-CHAIR TIRSCHWELL: Okay. Any other discussion on this or should we go ahead. Michael, do you have any closing comments about all that? Summarize that for us and help us --

MEMBER KAPLITT: Well, if you do them-- I mean, if you do the math, unless the developer can, you know, show me different numbers, if you do the math, all of the other exclusions the numbers make sense actually though. That's what I was doing there. The calculations here they

all make sense relative to that two million two hundred thousand denominator, right.

The only one that doesn't is the reason for not initiating IV therapy. If you take that number against that denominator it's only three and half percent. So, if it's a simple arithmetic error then it's not a major concern going from one percent in the original group to three and half percent and now excluding I wouldn't be too worked up about it personally. But going from 1 percent to 25 percent to 24 percent, you know, that's a different thing.

So, you know, that's the whole thing. I assume it's a simple arithmetic error in somebody's part to put those numbers here. Then fine, I don't see a major validity. All the other issues notwithstanding, clearly they're not affecting it that much. It's only gone from 1 to 3-1/2 percent. But if it's gone from 1 to 24 percent then all these things everybody has been raising could matter.

CO-CHAIR TIRSCHWELL: Yes, I mean on a

personal level I feel pretty confident that there is the math error and it's not nearly as big a change.

MEMBER KAPLITT: I do too.

CO-CHAIR TIRSCHWELL: Because I think the denominator probably got changed in that one but you're right. We can't be 100 percent sure about that.

MR. SCHMALTZ: Yes, this is Stephen Schmaltz from the Joint Commission. I do believe it is math error.

CO-CHAIR TIRSCHWELL: All right. So, let's go ahead and --

CO-CHAIR KNOWLTON: Well, I have a question.

CO-CHAIR TIRSCHWELL: Yes, sorry, go ahead.

CO-CHAIR KNOWLTON: So, what is the recommendation of your work group? I'm concerned that we're supposed to -- maybe I'm being too parochial but we're supposed to vote about what's in front of us. Now if it's a math error, it's a

math error, but what's in front of us and what's the recommendation.

MEMBER KAPLITT: Well, the original recommendation was -- I think it was like moderate or it was acceptable validity. But this issue hadn't been raised. This was something that came up when I was preparing for this where I noticed that we hadn't appreciated this and didn't discuss it. So, that's why.

So, the original recommendation based on everything else was that it was acceptable whatever, you know, validity. If it was a simple math error I think the gist of our call and our work group call would hold in that regard, you know. If it's not then everything changes.

CO-CHAIR TIRSCHWELL: Reuvin.

MEMBER FERZIGER: Peter's question. Did I understand that the exclusions may be different according to which kinds of hospitals they're coming from?

CO-CHAIR TIRSCHWELL: I mean the exclusions is kind of a safety net so that if you

feel like you -- you're the clinician, you feel like you have a valid reason why you're not giving tPA then this is not going to count against you that you don't give it. You just have to document that valid reason and different providers may feel that they have a different list of valid reasons. And that's all acceptable.

MEMBER FERZIGER: Did I understand you correctly that the rates of exclusion and the kinds of the rates were different based on the subjects that they're coming from?

DR. SCHWAMM: I would make two general comments. The first is that the Joint Commission -- so, I wear many different hats in this field, but at the moment I'm wearing my Joint Commission hat. But as my American Heart Association get with the guidelines measure developer, we have a slightly different approach which is we actually give a list of what we consider acceptable reasons based on the AHA guidelines, and we revise that every time that AHA changes that

list.

There is, I think, room for subjective interpretation about certain characteristics whether a patient's appropriate for treatment or not. All of the measures have reasons for non-treatment that are considered appropriate or valid. The Joint Commission standard is to require that somebody document in the record that there was a specific reason for not treating that link to the treatment. So, I didn't treat him because he was 93, and I felt his risk of hemorrhage was too high. They just collect that the reason was documented, and they actually collect reports and check off a box that says what was the reason. And they have the ability to write in a reason but if it's not on that list it doesn't.

I think either approach is very reasonable and valid. The more granular approach is more burdensome and onerous and so I think it's a very reasonable approach to say you must document the reason why and then rely on the

hospitals to review their cases and look for patterns of systematic bias or discrimination where they're not treating certain racial groups, ethnic groups, gender or just not treating anybody.

One thing you can do is look at the actual numbers of denominators and see if the denominator is shrinking and that's why rates are rising or if, in fact, the rate is going up because the enumerator is actually going up even as the denominator goes up. So, I think in this circumstance I do think it's very reasonable to accept a documented exclusion. and I think every measure must have that, otherwise it's like no risk adjustment. Otherwise, if you have a lot of people are inappropriate for the measure you somehow penalized, then you have a perverse incentive to treat everybody.

I think it is possible that some hospitals use the exclusion more frequently than others, but overall the rates of tPA use have been rising pretty dramatically across just

absolute rates. Used to be 3 percent; now they're closer to 8 percent. So, there has been over the last decade a continuous significant linear increase in the use of tPA.

MEMBER FERZIGER: It all makes sense; it just raises the spectrum of why the exclusions are being applied differently to different groups from different centers, and that would be a disparity of some concern.

DR. SCHWAMM: Agreed.

CO-CHAIR TIRSCHWELL: Okay. Let's go ahead and vote on validity.

MS. OGUNGBEMI: We are not voting on validity for Measure 0437. The options are high, moderate, low and insufficient. Voting is open.

CO-CHAIR TIRSCHWELL: Sorry, did you have another question, Valerie? I apologize.

MEMBER COTTER: Could you give us your advice on voting for this particular point?

CO-CHAIR TIRSCHWELL: Honestly, you know, I do believe this math error thing so I'm not worried about the validity here, and so I'll

be voting accordingly.

UNIDENTIFIED PERSON: Are we open?

MS. OGUNGBEMI: Yes, we are open.

Voting is closed. The results are 4 percent high, 65 percent moderate, 17 percent low and 13 percent insufficient. Measure 0437 passes on validity.

CO-CHAIR TIRSCHWELL: Feasibility, Mike?

MEMBER KAPLITT: Yes, I'm just looking at my notes here. I just wrote that most of the data elements in which were generated, and I don't think there are any major feasibility issues assuming nothing new has come up since the last time.

CO-CHAIR TIRSCHWELL: Any discussion? Let's go ahead and move to vote on feasibility.

MS. OGUNGBEMI: Voting on feasibility for Measure 0437 is open. Options are high, moderate, low and insufficient.

Voting is closed. Results are 43 percent high, 57 percent moderate, 0 percent low

and 0 percent insufficient. Measure 0437 passes on feasibility.

CO-CHAIR TIRSCHWELL: And then usability and use?

MEMBER KAPLITT: Usability is the same thing. I mean, I don't have a whole lot to say. It's, you know, I think that, you know, the benefits outweigh obviously I think most of the unintended consequences. Under usability, I'd put the comment that you had made earlier about this issue of, you know, physicians being able to put any measure -- you know, any medical reason that they want, but I don't think that's a major issue.

CO-CHAIR TIRSCHWELL: Discussion? Go ahead and move to vote then on usability and use.

MS. OGUNGBEMI: Usability and use is what we are now voting on for Measure 0437. Voting is open. The options are high, moderate, low and insufficient.

Voting is closed. Results are 52 percent high, 48 percent moderate, 0 percent low

and 0 percent insufficient. Measure 0437 passes on Usability and Use.

CO-CHAIR TIRSCHWELL: So, then I think we just have the general suitability for endorsement vote. Discussion before that, Charlotte?

MEMBER JONES: Do we not need to discuss unintended consequences?

CO-CHAIR TIRSCHWELL: I think that should have been done already in the Usability and Use. So, we kind of missed that. Do you want to feel free to raise your point and we can decide whether we need to do something different?

MEMBER JONES: Well, I think that as Stephen pointed out with the previous one -- and it has been reported in the literature -- we know that hospitals that work on improving their time to needle increased treatment of minutes -- of stroke mimics and that's in the evidence. And I think it should be addressed. I think the fact that it wasn't even mentioned that there was a published article in a well-respected journal is

concerning to me and raises the question of transparency.

CO-CHAIR TIRSCHWELL:

MEMBER RAE-GRANT: This was raised at the last time. Steve raised that article and discussed that the last time. I guess we would just have discussion about that mimic issue.

CO-CHAIR TIRSCHWELL: That was --

MEMBER RAE-GRANT: It's on record as part of the discussion.

CO-CHAIR TIRSCHWELL: For the different measure, you mean?

MEMBER RAE-GRANT: Different measure, but at least we had discussed it in some way at committee. I'm pointing that out, yes.

MEMBER JONES: And that may have been a discussion that I wasn't part of. I was part of this one and I am, in fact, the person who sent it in into the group that it should be discussed. I think we need to discuss it as a group that at least we raised this issue as an unintended consequence of this measure.

CO-CHAIR TIRSCHWELL: Okay. I hear that you have raised it, and I guess I would ask: does anybody else have any comments or discussion about the unintended consequence of pushing hard for treatment leading to use of more tPA in false/positive stroke looking patients? Anybody on the committee? Dr. Schwamm?

DR. SCHWAMM: Yes, the only comment I would make is that it's been well demonstrated in the literature that the risk of harm to patients who have stroke demonstrated with tPA is very, very low, less than .5 percent. And it's also been demonstrated that the exponential nature of the decreasing effectiveness of the drug as time goes on, and I would argue pretty strongly the population attributable benefit to those who get treated by not waiting greatly outweighs the risk of harm to those who are treated rapidly but who turn out later on to have been a mimic. In many of those patients, it's not possible to determine their mimic status in a rapid manner. It's a judgment call. And we do lots of -- we do

appendectomies on patients that we think might be having appendicitis, knowing that 30 percent may not actually have had one because we know the benefits are so striking in the treated population. So, I would just argue I think that rapid treatment is still justified.

CO-CHAIR TIRSCHWELL: Steve then Charlotte.

MEMBER HUFF: Everybody is in favor of rapid treatment and we hope to get more accurate, more specific biomarkers in the future. I think currently with regards to appendicitis, no one has that false negative rate because everyone gets imaged for -- virtually everyone gets imaged first now.

I just think that an unintended consequence of pushing is -- and it's interesting. The papers that say there's very few adverse reactions for the stroke mimics getting tPA are somewhat limited in number, and back in the days when we used to give tPA intravenously for cardiac conditions more

frequently, which is a much lower dose of tPA, there certainly was a recognizable complication rate so it doesn't make sense to me that we're giving more tPA and yet we seem to have no complications on a similar patient population. So, I'm just concerned about an unintended consequence, and it would be nice to monitor that in some way. I realize that may not be germane to this discussion.

CO-CHAIR TIRSCHWELL: Right. Charlotte, go ahead.

MEMBER JONES: I am going to again as a pediatric neurologist, this measure goes down to the age of 18. It may very well be true that if you are talking in the 35, 40, 55 and up population that the risk of treating a mimic is reasonable on the number needed to treat. But if you were giving 18 and 19 year olds --- and I realize I don't have evidence either, but I think that their National Quality Forums should request -- and I think that even in the discussion that knowing that people are publishing the increasing

time to needle, increases treatments in mimics -- that we have a responsibility to at least ask developers to track that. So the 18, 19 and 20 year olds who are coming in with their complicated migraines or their post concussive symptoms and they're being treated because we are saying you have this period of time so you may not be able to get the history and find out that this was, in fact, a Todd's Paralysis, which is much more common in the younger population than the older population, that we have that evidence because you're going down to 18. And the risk/benefit and the positive pre-test probability are just very low in the 18 and 19 year olds.

CO-CHAIR TIRSCHWELL: Thank you for those comments, Charlotte, and this will be absolutely noted in the reports and the summary from the committee. And I, unless somebody objects, I'd like to go ahead and suggest that we just revote on Use and Usability taking into account these new comments. Does anybody object

to revoting?

Michael, go ahead.

MEMBER KAPLITT: And I would simply point out that when we were questioning before about the value of the imagine measure, this is one of the values which is to reduce the unintended consequences of giving something to someone who could be harmed by it. So, you know, that's why we have the other measure to try to protect people.

CO-CHAIR TIRSCHWELL: Reuven?

MEMBER FERZIGER: Just a question to Charlotte. And that is, now I understand the risk of unintended consequences in younger patients may be significant, maybe. But how common are they, you know, as a chief complaint something that would lead to tPA in the ERs?

MEMBER JONES: So, what I can tell you right now is in one large children's hospital we have been asked to develop a stroke team, and we have been tracking our data, and so far we are seeing -- we've seen 12 complicated migraines

that were called by the ED as an acute stroke. None of them were. And we have had issues with mis-reads of MRIs in children with shunts who had Todd's Paralysis because the shunt impacts on the quick MRI that ordinarily the radiologist want to do and they've read ischemic lesions which was shunt artifact -- all of which means that as we push a younger and younger age group, there be risk benefit. It's like the adult cardiologist who says to me, well, gosh, if this was 50-year old, I'd tell you that 3-year old was having an acute MI. I have to say well, the pre-test probability of an acute MI in a 3-year old is pretty damn low.

MEMBER FERZIGER: So, isn't it a reasonable question to measure development to ask if this causes you to rethink the age criteria for this measure or not?

MS. KOLBUSZ: For the Joint Commission all of our in-patient hospital measures are for the adults in patient population, so under 18 are excluded. I think that those are all very

important facts, and we recognize that there is a pediatric population but the measure does not address the pediatric population whatsoever.

CO-CHAIR TIRSCHWELL: Yes, we realize that the 18-year olds are not the same as 70-year olds. Is that what you were going to say, Charlotte?

MEMBER JONES: I was going to say the American Academy of Pediatrics says they're pediatric until they're 21.

CO-CHAIR TIRSCHWELL: And I think Obamacare puts it up to 26. They can still be on your insurance, right?

UNIDENTIFIED PERSON: We will not solve the debate on when your child is no longer a child.

CO-CHAIR TIRSCHWELL: Any other comments or discussion before we go ahead and revote on Usability and Use? Let's go ahead then.

MS. OGUNGBEMI: Voting on Usability and Use for Measure 0437. This is a revote. The

options are high, moderate, low and insufficient. Voting is open.

Voting is closed. The results are 30 percent high, 65 percent moderate, 0 percent low and 4 percent insufficient. Usability and Use it passes Measure 0437.

CO-CHAIR TIRSCHWELL: Great. Now we'll go to overall. Any discussion -- further discussion before the vote on the overall measure?

So, we're open for Overall Suitability for Endorsement.

MS. OGUNGBEMI: We are now voting on Measure 0437's Overall Suitability for NQF Endorsement. Options are yes and no. Voting is open.

Voting is closed. Results are in. Unanimous 100 percent votes, yes. Measure 0437 is suitable for NQF endorsement.

CO-CHAIR TIRSCHWELL: Great. So, that means that we will also be going through the eMeasure 2834. Did you guys want to add any

comments about the eMeasure before we go in or did you already comment about it?

MS. KOLBUSZ: You know as far as the evidence and all, it's the same.

CO-CHAIR TIRSCHWELL: Right.

MS. KOLBUSZ: I think you established at the beginning of the meeting that as far as the eCQMs are concerned, they haven't been validated. We have used the Bonnie Tool for testing information. And I would actually turn over the discussion, I think, to our eCQM nurse informatics person because she could be -- answer specifics about the eCQM.

CO-CHAIR TIRSCHWELL: Just a brief overview is okay.

MS. ANDERSON: Sure.

CO-CHAIR TIRSCHWELL: Thanks.

MS. ANDERSON: Oops, sorry. I have to get my computer opened up.

Okay. So, as Karen said the eMeasure mimics the Chart Obstructed Measure very closely. Our description is the same description as the

Chart Obstructed. We are measuring for acute ischemic stroke patients who arrive at this hospital within two hours of known well and for whom tPA was initiated at this hospital within three hours of the last time well.

Our denominator is looking at ischemic stroke patients admitted to the emergency department whose arrival time is within two hours or less. Less than or equal to 120 minutes of the time they were known to be at their baseline state of health or time of symptom onset is less known at the time is not known.

Our denominator exceptions are excluding patients with comfort measures documented on the date of or date after arrival. Patients with IV or IA thrombolytic therapy prior to arrival, patients with the documentation of a NIS score of zero in the emergency department, patients with medical reasons for not initiating IV thrombolytics documents by a physician, APN, PA or pharmacist on the day of or day after arrival.

Patients with the following results within 180 minutes of the time they were known to be at their baseline state of health or time of symptom onset. These include prothrombin time greater than 15 seconds, platelet count less an 100,000, INR greater than 1.7, partial thrombol less time greater than 40 seconds, systolic blood pressure greater than 185, and diastolic blood pressure greater than 110 and patient refusal.

You can tell we get a little bit more granular in our data elements with electronic clinical quality measure than what we have on the chart obstructed measure.

The numerator, we are looking at acute ischemic stroke patients for whom IV-tPA was initiated at this hospital within three hours. Less than or equal to 180 minutes of when it was witnessed or reported that the patient was last known to be without the signs and symptoms of current stroke or his or her baseline state.

This measure has been adopted by the EHR Incentive Program and the hospital and

patient quality reporting program as an electronic clinical quality measure --- CMS-91.

The current format of this measure was first specified in December of 2012, and it has been used by hospitals attesting to meaningful use program.

When reporting this measure, hospitals attested to the eCQM specifications, approved by CMS for use at the time of the reporting period, thus indicating feasibility of the measure.

In addition to the meaningful use program, the measure of feasibility is supported by the fact that this measure was used by hospitals voluntarily submitted eCQMs to the hospital and patient quality reporting program in 2015.

In 2016, CMS is requiring organizations participating in the HIQR program to electronically submit one-quarter of data for the 28 available eCQMs, and this measure is one of those measures that they can select from. However, we do not have any data currently on the

eMeasure.

CO-CHAIR TIRSCHWELL: Okay. Mike, is it you again or --

MEMBER KAPLITT: Yes, I guess. So, do we really need to re-review the evidence on the gap? We just did that, right?

CO-CHAIR TIRSCHWELL: I don't think so.

MEMBER KAPLITT: Right. So, we can get right to reliability I assume?

CO-CHAIR TIRSCHWELL: I think that's correct and acceptable so, yes.

MEMBER KAPLITT: Okay. So, with respect to reliability there's one major issue which is as an eMeasure, our understanding we discussed this on the work group is that NQF, you know, the NQF standards say that the eMeasure has to have been shown to be -- has to have been tested in electronic health records for more than one vendor, and there is no evidence of that here. And that's a legitimate concern because obviously we all know that for an eMeasure, the reliability is very much based on, you know, who

the vendor is and whether you're actually going to be able to capture what you want to capture in a reliable way.

CO-CHAIR TIRSCHWELL: So, it seems like there's been a shift on this issue of having already been tested in the EHR for approval. Do you guys want to comment on that? I mean, clearly if we're thinking about approving this measure and there are those -- there are no such data about the EHR then we're not requiring that to get to approval. Is that correct?

MS. JOHNSON: This is one of those funny measures that we're calling legacy measures.

CO-CHAIR TIRSCHWELL: Yes.

MS. JOHNSON: Correct? So, they are already in use in Federal programs, so what we at NQF have done is we're allowing use of this, I think Ann calls it simulated data, to kind of stand in for testing. We only do that for these legacy eMeasures.

MEMBER KAPLITT: Well, then I don't

know what to add. I mean, I personally disagree with that, but I don't know what to add. I mean, because I think the Bonnie Test relates to validity and that I'm fine with, but I don't think it relates to reliability. But that's fine if that's the new standard. So I don't have a whole lot to add. I don't know if anybody else wants to but I just, you know, for me I don't see how we can pass -- and I don't understand a legacy standard is different than any other standard because we're going to hold people to a standard that hasn't been tested in a sufficient number of VHRs to show that it's actually reliable and usable, I don't why it should matter whether it's a legacy measure or not. But, you know, maybe I'm in the minority.

CO-CHAIR TIRSCHWELL: And you said this is being used already by CMS?

MS. ANDERSON: Yes, it's currently in use.

CO-CHAIR TIRSCHWELL: Since when?

MS. ANDERSON: Since 2012.

CO-CHAIR TIRSCHWELL: Since 2012? So, how is it that we have no data?

DR. SCHWAMM: Reintroducing Ann Watt.

MS. WATT: CMS made these measures part of the meaningful use program without them ever actually having been implemented, and the requirement was that hospitals had say that, yes, we can collect the data on these measures. So, they've been in the program for that period of time. Actual data collection has not been required yet; it is just now in 2016 being required, as Lisa explained, for what the third or fourth quarter needs to be reported to CMS by the first quarter -- by February of 2017. So, that's why we have no data. These measures have been around since -- for a very long time, but data collection has not been required until just now beginning in 2016.

CO-CHAIR TIRSCHWELL: So, then what was the rationale for bringing it forward for approval at this point before there was the data?

MS. WATT: We were requested to do so,

and we were given the guidance that Bonnie Testing would be sufficient to attest to the reliability of the measure.

CO-CHAIR TIRSCHWELL: Okay. Any other -- yes, go ahead, Mike.

MEMBER KAPLITT: I can just tell you that just because CMS said something a few years ago, those of us who have been struggling with the meaningful use ever since that day can attest that it, you know, a lot of these EHRs have not kept up, that we've spent the last four years or whatever it is constantly trying to modify and adjust things to try to capture that because it's not always been sufficient. So, I'm not sure just because CMS said we should do it that that means that eMeasure is ready for prime time across the board. That's all.

CO-CHAIR TIRSCHWELL: Any other discussion before we go ahead on vote on reliability? Any other comments from the developer?

MS. WATT: No, my only comment be that

we provided the information that we were requested to provide for this measure.

CO-CHAIR TIRSCHWELL: Yes, Reuven, go ahead.

MEMBER FERZIGER: We're back on the policy domain, so I really would be quite interested given the amount of thought that Mike has given into this and the other perspective of the Chairs on advice about voting.

CO-CHAIR TIRSCHWELL: Well, you know, so I certainly didn't make the rule that Bonnie Testing would be adequate for demonstrating reliability. In fact, when we first talked about these measures in our working group this was -- this was rated as insufficient, specifically because there was no EHR data. And then there's been a political shift to alter that perspective on the whole thing. That was done completely outside of the purview of our committee, and I don't -- you know, I don't even begin to believe that my pay grade is anywhere close to being able to comment on that.

Honestly, I'm a little uncomfortable with it but -- and you can vote your conscience, even if it is outside of standard algorithms. That is within the purview of this committee, if I'm correct. So, I guess I don't have much more to say about it than that.

MS. JOHNSON: So, let me put a little bit of context around this, and I do realize this is confusing.

One of the things that NQF allows is what we would call -- if a developer demonstrates data element validity, then we do not require additional reliability testing. That's been something that has been the case for many years now. So, that's one thing.

This measure originally came through and they had demonstrated element validity through the Bonnie Testing tool, which we also said was appropriate. What the developers had not done when you had looked at it in the work group is they had not addressed the threats to validity as things were blank. And that is why

we, as staff, originally selected insufficient as our rating.

Since then, the developer did add in some information on the threats to the validity, so that's how it has moved from insufficient to something else.

So, what you're looking for reliability in the case that a developer does not do separate reliability testing but relies on their data element validity testing, you have to see what you would rank or how you would rate their data element validity testing and just use that rating in reliability. And you'll see how that works in the algorithm. So, I do realize it is confusing. And I apologize for that, but that's the background of going from insufficient.

CO-CHAIR TIRSCHWELL: Ron?

MEMBER KOENIG: Would you please clarify one thing for me? In the previous state policy we had, the numerator and denominator were quite clear. The numerator were those who received, and the denominator were those who were

eligible. In this one, the denominator includes time of symptom onset -- the time last known at baseline rate is not known. What does that mean?

CO-CHAIR TIRSCHWELL: Does the developer want to comment on that?

MS. ANDERSON: So, your question was about why in the eCQM do we have data elements for specific time of symptom onset and baseline--

MEMBER KOENIG: In the first one, we had those who received and those who were eligible. On this measure, we have those who have received, but the denominator is anyone who has a stroke whether they're eligible or not so it seems by the wording.

CO-CHAIR TIRSCHWELL: No, I mean they still have to be within the time frame and there's a number of other exclusions.

MEMBER KOENIG: Well, the time of symptom onset, the time last known of baseline state is not known. So, what's the time?

CO-CHAIR TIRSCHWELL: If the denominator, sorry. Principal diagnosis of

ischemic stroke or -- yes. I think maybe the denominator details which are in the documents and I guess ischemic stroke patients within two hours -- yes. Ischemic stroke patients who present within two hours. And then there are exclusions on top of that that can --

MEMBER KOENIG: Does that mean if someone went to sleep and sometime during the night had a stroke and then the family brings them in late in the morning or --

CO-CHAIR TIRSCHWELL: No, they're out.

MEMBER KOENIG: Okay.

CO-CHAIR TIRSCHWELL: It's last known normal.

MEMBER KOENIG: Okay.

MS. ANDERSON: Right, so the baseline state and the time of symptom onset are reflective of that -- the old data element of last known well. So, in order to specify it a little bit better in the EHR, we have two data elements that when you look at the logic of the eCQM, they are worse statements, so we were

either looking for the baseline state or the time of symptom onset.

CO-CHAIR TIRSCHWELL: Yes, it's confusing.

MS. ANDERSON: It's a little bit confusing.

CO-CHAIR TIRSCHWELL: The way it's worded by --

MS. ANDERSON: In the way it's worded and the way it's specified.

CO-CHAIR TIRSCHWELL: I think the intent makes sense.

Sorry, so let's see. Where were we? So, Karen, just to summarize what you said, the change from insufficient to more acceptable was based additional information from the developers about threat to exclusion. It was never about the lack of data from EHRs. Is that an accurate statement?

MS. JOHNSON: I believe I would say that is accurate. We have told developers that we will accept Bonnie Testing information. So,

for the testing, that is what we have. We couldn't accept not giving anything on the threats to validity. So, they did come back and tell you something about exclusions, and I believe they maybe also included some information about meaningful differences.

And if you would scroll down, Alexander, to the validity section, you can see that we did provide you and we red-lined it so that you would know that we went in a little bit later and made changes. So, yes, so I'm remembering correctly. They did add some information about exclusions to meaningful differences.

CO-CHAIR TIRSCHWELL: And the line in the policy that Mike referred to about testing in two EHRs, that line is really in there, but it refers to a new electronic measure that's not a legacy measure? Is that an accurate statement?

MS. JOHNSON: It refers to a non-legacy measure. I have to think about whether it's only the new ones. Depends on what you mean by new,

but I think it's correct to say --

CO-CHAIR TIRSCHWELL: Well, it's an electronic measure that's not a legacy measure.

MS. JOHNSON: Right.

CO-CHAIR TIRSCHWELL: And we don't have any of those, right?

MS. JOHNSON: There are some. Because electronic measures that -- what we mean by legacy is that they are being used in Federal programs such as meaningful use. So, we carved out a subset of eMeasures, and we are allowing Bonnie Testing for that subset of eMeasures. So, the one that you have in front of you now is one of these legacy eMeasures. It's being used in the meaningful use program.

CO-CHAIR TIRSCHWELL: But we're not -- neither today nor tomorrow, are we reviewing an eMeasure that isn't a legacy eMeasure?

MEMBER KOENIG: I think the --

CO-CHAIR TIRSCHWELL: Or hybrid maybe, that's different.

MS. JOHNSON: The hybrid measure is one

of them.

CO-CHAIR TIRSCHWELL: That field is very different so -- and they have a lot of data.

MS. JOHNSON: Yes, and you also have a couple --

MS. MUNTHALI: Karen?

MS. JOHNSON: Yes.

MS. MUNTHALI: We're also going to look at approval for trial use --

MS. JOHNSON: yes.

MS. MUNTHALI: -- which is, I believe, hey have Bonnie for that as well.

CO-CHAIR TIRSCHWELL: That's different though also.

MS. JOHNSON: Yes. So, no, I mean other than the hybrid eMeasure, we're not. But other --

CO-CHAIR TIRSCHWELL: Okay.

MS. JOHNSON: -- other projects have.

CO-CHAIR TIRSCHWELL: All right.

So, thank you for clarifying. Any other questions? Charlotte, go ahead.

MEMBER JONES: So, I think this is probably the equivalent of a math error, but are you excluding patients under the age of 18 for this measure?

MS. ANDERSON: Correct. Our age requirement is greater than or equal to 18 years.

MEMBER JONES: Okay. It's not documented here.

MS. ANDERSON: It's in our initial patient population, so it is included. It's just not on this document.

MEMBER JONES: Okay. Thank you.

MS. ANDERSON: You're welcome.

CO-CHAIR TIRSCHWELL: Okay. Peter, go ahead.

MEMBER SCHMIDT: So, listening to the description of where we are, the reliability switched from being insufficient to being something else because of the -- because of your assessments and things like that. But I still see that the guidance from the algorithm says low.

MS. JOHNSON: That was our mistake.

MEMBER SCHMIDT: So, what would the guidance from the algorithm be now, just so that we have just a sense of what the recommendation is?

MS. JOHNSON: The guidance from the algorithm for validity we would say right now is moderate?

MEMBER SCHMIDT: For reliability?

MS. JOHNSON: For reliability, since they don't have other testing data for reliability, it would take on the rating that it would get for data element validity.

MEMBER SCHMIDT: Okay. So, but if we vote low on reliability, are we going to get to valid -- to validity?

MS. JOHNSON: No. It would not pass.

MEMBER SCHMIDT: So, validity would trump reliability, but reliability tails validity?

MS. JOHNSON: You know --

MEMBER SCHMIDT: There's a movie called

Catch 22, I think.

MS. JOHNSON: Yes. It is tricky because, you know, a reliable measure, you know, by definition needs to be valid but also valid measures we want to be reliable. So, we have to pick one.

MEMBER SCHMIDT: I consider myself tricked.

CO-CHAIR TIRSCHWELL: We should vote on validity first.

Okay. So, Charlotte, do you have another comment and then Alex, go ahead.

MEMBER RAE-GRANT: Just in terms of time management, we have four more guidelines to do in the next five minutes, just to remind us where we are.

CO-CHAIR TIRSCHWELL: Yes. We fell off of that one, thanks.

All right. Let's go ahead and vote on reliability.

MS. OGUNGBEMI: We are now voting for Reliability on Measure 2834. The options are

high, moderate, low and insufficient. Voting is open.

Voting is closed. Results are zero percent high, 43 percent moderate, 35 percent low and 22 percent insufficient. We are landing in a gray zone.

CO-CHAIR TIRSCHWELL: I mean, a gray zone suggests to me that we probably need to go through the rest of the criteria, and then at the end the overall vote will sort of deal with this gray zone here.

Am I running afoul of any policies at NQF in that suggestion? Does it seem okay with you guys?

MS. JOHNSON: I think I want Elisa and Marsha to handle that one.

MS. MUNTHALI: So, you would vote on overall suitability for endorsement, but we will consider the gray zone, consensus not reached issues in the first consult.

CO-CHAIR TIRSCHWELL: Okay. So, validity? Mike?

MEMBER KAPLITT: We already discussed this. They did the Bonnie Testing, the validity was good by 100 percent of the various, you know, they did like 23, I think it was, patient -- synthetic patient charts so everything was covered and validity testing is not the problem here.

CO-CHAIR TIRSCHWELL: Discussion on validity? Sorry, did I cut you off? Anything else?

Let's go ahead and move to voting on validity.

MS. OGUNGBEMI: We're now voting on validity for Measure 2834. The options are high, moderate, low and insufficient. Voting is open.

Voting is closed. The results are 17 percent high, 61 percent moderate, 9 percent low and 13 percent insufficient. The measure passes on validity.

CO-CHAIR TIRSCHWELL: Feasibility?

MEMBER KAPLITT: I don't know what to say. I mean, you know, I don't think that the

issue is any different, right, in the absence of as an eMeasure -- not as a measure. As an eMeasure, I just don't see how we can comment on feasibility without real-world data in the actual records that are used. But -- I mean, I think the measure itself, we talked about, an hour ago, we said that was feasible but I don't see how we can comment on feasibility.

MS. JOHNSON: I'm just curious, was there a feasibility score card filled out for this one?

MS. ANDERSON: No. It would be feasibility ---

MS. JOHNSON: We didn't have a score card; we did do a feasibility report in lieu of the score card.

MS. JOHNSON: Yes, the score card came before these measures or came after these measures were actually in use. So, there were no real feasibility scoring done the way that a brand new measure would be, so we did do a feasibility report which should have been

included in the documentation.

CO-CHAIR TIRSCHWELL: Reuven?

MEMBER FERZIGER: So, therefore, how would it be possible to vote on anything other than insufficient for feasibility, since we have neither the report nor any information about real-world feasibility?

CO-CHAIR TIRSCHWELL: So, I would defer that question to NQF staff. What are we voting on if we don't have any evidence of feasibility in an actual EHR? Is it just that the logic and required data are readily available without undue burden? Because they can be implemented for a performance measurement part at the end there would seem to be conjecture, I guess at this point.

MS. JOHNSON: Ann is going to address that. Can you turn on your mic, Ann?

MS. PHILLIPS: Yes, feasibility assessment was based on Bonnie performance, so it shows that the measured logic is measured logic is functional. You are correct. It does not

show the measured maps in the EHR. So, we would believe all the data elements are associated with value sets. All the value sets are published and the lead back and should be commonly found in the EHR.

CO-CHAIR TIRSCHWELL: So, is evidence that it can be implemented in a real EHR required for feasibility?

MS. PHILLIPS: Not for a legacy measure.

CO-CHAIR TIRSCHWELL: Okay, yes, Charlotte, go ahead.

MEMBER JONES: I just pulled up the data accuracy feasibility report, and for those people who don't have it I'll just read it. "At this time, we are unable to directly assess the accuracy of these elements in an EHR system. However, because these data elements are used across multiple measures and are harmonized with the chart instructive version of the measures with which hospitals are already familiar, they are likely to be monitored closely for

correctness. In addition, Data Element 14 was created based on feedback at the request of implementers to improve feasibility and is believed to be highly feasible."

CO-CHAIR TIRSCHWELL: Oh, sorry, go ahead, Peter.

MEMBER SCHMIDT: I'm wondering if this was included in meaningful use. Doesn't that mean somebody has collected information about this and we should be able to say that it was feasible or is nobody doing meaningful use?

MS. ANDERSON: People are reporting meaningful use, but they've been attesting that they can collect this data which is a little different than actually being able to extract the data the way that we would for an eCQM. I mean, it's a proxy, right, but --

CO-CHAIR TIRSCHWELL: Yes, they've said consistently that the data are not available at this time.

Sorry, Charlotte, are you still up there? Steve, you're thinking about it I can

tell.

MEMBER HUFF: Yes, I mean, as a non-statistician, non-IT person it's just hard for me to know, taking 31 synthetic charts and extrapolating that to tens of thousands of real charts, how real that is. So, I really depend on guidance. I have no professional opinion on this.

MS. WATT: Let me just give a little bit more of information about what Bonnie testing is. And speak up if I'm mis-speaking. But basically, test cases are created in the Bonnie System. That's actually what it is, is it's testing bed that reflects all of the possible answers to all the possible data elements in a measure. Those that would exclude it as well as those that would include it. And where the 31 cases came up with is, that's the total number of permutations of the data element and the allowable value that would be able to say that the fair degree of confidence, yes. This measure logic computes the way we expect it to compute.

That actually is I think a pretty good proxy for feasibility because at least we know the measure logic works in all possible permutations of the data.

DR. SCHWAMM: And I think when you think about feasibility with an eMeasure, what you want to know is when you push the button, do all of the inclusions and exclusions work together in the flow of the measure construct and kick out an answer? It doesn't tell you whether the correct values were populated into the EHR, but it does show you that when you push the button you get the report. And so I think that just means that the software is valid when tested against a sample data set. It does not tell you the quality of the data entered into those value fields, but they were the correct formats. Dates were dates, times were times. Yes, no's were yes, no, etcetera.

CO-CHAIR TIRSCHWELL: I guess how does it -- how is it different from validity then at that point? It seems like the Bonnie testing

really squarely hits on the validity thing. I guess I don't -- it seems like feasibility should not be the same as validity. And --

DR. SCHWAMM: I'll go out on a limb and maybe be abandoned by the Joint Commission when I say this, but it seems to me the measure if valid that if in this case there's a human factor involved as well which is instead of the abstractor looking at the whole record and trying to come up with the best answer, you're relying on whoever entered that data into the EHR field at the time during the process of care that they did. And so if a medical student writes down that the time of stroke last seen well was noon, but the attending comes by later and writes down actually it was 4:00 p.m., but noon is what populated that field, then the measure construct is feasible. The construct is valid, but the data that's been entered may or may not produce an accurate report on that specific patient. I'm not sure which bucket that falls into. Reliability, I guess.

CO-CHAIR TIRSCHWELL: I guess what happens if we roll this out and then the reality is that it is an undue burden for these data to be collected, and they are very inconsistently completely reported then after the fact we would say, oops. I guess we were wrong about feasibility, which maybe is okay. But it seems like that's to me just the seat of my pants, that's what feasibility is about.

MS. WATT: I was going to that this is Ann from the Joint Commission, but we know that.

You know, the thing about these measures particularly is that they have been in existence for a long time. We know that a lot of EHRs, all of the EHR systems are collecting data on these measures. We know that CMS has collected data on these measures in a testing mode, and we know that the Joint Commission has had. And we also know that based on the results -- you know, this has sort of been an iterative process, and based on what we have learned changes have been made to the measure constructs

to make them feasible and collectible.

Every hospital who is reporting meaningful use has said, yes. This is a feasible measure for us to collect, and we can do it.

CO-CHAIR TIRSCHWELL: But what you just said is that the data have been being collected and that you've learned something from them, but you didn't share any of the data that you --

MS. WATT: We don't have access to CMS data and we -- the volume of data that the Joint Commission received during our test period was not sufficient to be able to do that. We have not, you know, put this measure into production on a wide scale.

CO-CHAIR TIRSCHWELL: Comments or questions? Yes, Reuven, go ahead.

MEMBER FERZIGER: So, I think that in this particular day keep running into this issue of, you know, what's policy and what's the scope of our job. And I think it may be excellent policy for this electronic measure to be deployed and for CMS to be collecting data on. And that's

above my capacity to judge. However, it seems that the scope of the job that we have is to judge whether feasibility has been established for this. I don't see any way in which feasibility, you know, by the definition applied anywhere else, you know, is available here. I think that it has to be insufficient. But I completely grant that despite the evaluation of insufficiency by the charge of this committee, it still may be a great idea to deploy it. I just don't see how we can say that feasibility has been established.

MS. OGUNGBEMI: We are now voting on feasibility for Measure 2834. The options are high, moderate, low and insufficient. Voting is open.

Voting is closed. The results are zero percent high, 26 percent moderate, 4 percent low and 70 percent insufficient. Measure 2834 does not pass on feasibility.

CO-CHAIR TIRSCHWELL: And it's a must pass criteria or --

MS. JOHNSON: No, feasibility is not a must pass criteria. So, you will continue your discussion.

CO-CHAIR TIRSCHWELL: All right. Moving along.

Usability and Use.

MEMBER KAPLITT: I mean, I don't have a huge amount to talk about. I mean, assuming everything else were met, then I think that Usability is fine. I mean, because I don't think it's that much different as an eMeasure than it was -- I don't think there's any data to argue that it's any different.

CO-CHAIR TIRSCHWELL: Any other discussion? Reuven, are you still up there? No problem.

Let's go ahead and vote then.

MS. OGUNGBEMI: We are now voting on Usability and Use for Measure 2834. The options are high, moderate, low and insufficient. Voting is open.

Results are in. Voting is closed.

Results are 9 percent high, 48 percent moderate, 17 percent low and 26 percent insufficient, and I believe we have landed in another gray zone. We landed in a gray zone.

CO-CHAIR TIRSCHWELL: So, we got two gray zones, one fail, and should we vote on overall endorsement? I think so. So, let's go ahead move to vote on overall endorsement.

And could you -- I think I'll do it this way. Could you just review the results of the -- since this has gone on for so long -- of the -- no discussion. I just want to know gray zone, pass or fail for the four criteria up until now.

MS. OGUNGBEMI: Yes, so for what we voted on, we have gray zone for Reliability. We have pass for Validity. Fail for Feasibility. Gray Zone, Usability and Use. So two gray zones, one pass, one fail.

CO-CHAIR TIRSCHWELL: Okay. So, you want some advice, Reuven? I don't know what to say. My personal feelings reflect the less than

overwhelming support for the different categories, and I'll be voting accordingly.

MS. OGUNGBEMI: We are now voting on the Overall Suitability for Endorsement on Measure 2834. Options are yes and no. Voting is open.

Voting is closed. The results are in and voting is closed.

We have 17 percent yes, 83 percent no.

CO-CHAIR TIRSCHWELL: I would definitely like to take a break, but I'm not sure we should.

So, let's see. You think we should? Okay.

MS. OGUNGBEMI: So, I have to say the measure is not passing on Suitability for Endorsement. So, no.

CO-CHAIR TIRSCHWELL: Thank you.

MEMBER BULSARA: Just a quick comment.

The developers are clear, I mean, it's not that our expectations are different from what the developers are sort of -- I mean, what their

expectations are, right? I mean there's a clear set of expectations ion terms of what would meet criteria in terms of passing these various things, because I just want to make sure that we're not judging on different criteria, like is there any sort of variation in terms of what the expectations are from them and in terms of what we're judging them for?

CO-CHAIR TIRSCHWELL: Well, you know, and you guys correct me if I get this wrong. They've heard the call to put this measure forward. The Joint Commission did, and they did their best to provide as compelling data as was possible. The National Quality Forum staff, you know, put it through the regular analysis and made some preliminary recommendations but if it was just algorithmic they wouldn't need us at all I would so. And so the point of us being here today is to bring, you know, the human factor to evaluating these criteria and, you know, lay that additional level of multi-stakeholder, common sense and evaluation to what has been performed.

So, I don't know. Do you guys want to add anything to that, or does that make sense?

MS. WATT: I makes complete sense. I would just like to emphasize that up until this morning the Joint Commission was under the understanding that all of these eCQMs were NQF endorsed. That apparently now is not the case. And so we brought them forward for re-endorsement because it is the three-year cycle.

This endorsement came about as an artifact of the eCQM development process, you know, going back to 2010. Determination was made somewhere along the line that if the source measure was endorsed, the eCQM was endorsed.

We have thought that these measures were endorsed. We brought them back for re-endorsement knowing that the data are not there. These measures have not been fully implemented yet. And we prepared the submissions according to the guidance of the NQF staff. Now, in fairness to everybody, I think that it's new for all of us. We appreciate the opportunity to be

able to discuss it. I have to tell you though that I sort of feel as though we have been -- the Joint Commission has been -- is being penalized for guidance that we received in good faith and followed in good faith, and thank you for your consideration. That's really all I have to say.

CO-CHAIR TIRSCHWELL: Break?

MS. OGUNGBEMI: I have one.

CO-CHAIR TIRSCHWELL: Yes, please.

MS. OGUNGBEMI: If no one else. Okay. So, we are going to have the committee dinner.

MS. MUNTHALI: I did want to address what Ann mentioned and any apologies if there's any confusion about the status of the eCQM measures that are under review today.

That was probably something that happened with the legacy measures. We have since changed our eMeasure process and policy. It's been about a year or two. We've been spending the last year or two trying to socialize this with developers and committees. And as you can understand, it's been very difficult to do

because it's not just dependent on what we do at NQF but also what is happening with the feasibility up to the completion of the measures but also with CMS and the availability of data. And so we'll continue to work with you and other developers to make sure that, you know, everyone understands what our policy is going forward. But this discussion has been very helpful for us because this is real life implementation of the policy and process. And so we're hearing some things about how we can refine it and we probably will be having further conversations with the Joint Commission and other developers and CMS to see how globally we can all improve this.

MS. OGUNGBEMI: Okay. So, we are having a committee dinner/Happy Hour if you all want to join us at Georgia Brown's just down the street, about a block away between the Metro and here. So, if anyone would like to join us after the meeting is over, could you please raise your hand? It would be around 6:15.

(Whereupon, the above-entitled matter

went off the record at 3:33 p.m. and resumed at 3:46 p.m.)

CO-CHAIR TIRSCHWELL: All right, we're going to go ahead and jump back in with the next set of companion measures, antithrombotic therapy by end of hospital day two. I guess I would invite you guys to just introduce them both briefly. And for all of us, we've gone through a bunch of issues here. We're happy to hear new information, but we probably don't need to rehash many of the same issues that may come up again with these other measures. Thanks, Karen.

MS. KOLBUSZ: Okay. The first is the chart-abstracted stroke five antithrombotic therapy by end of hospital day two. The eCQM does mimic or mirror the data elements and the construct of the chart-abstracted measure.

This measure captures the proportion of ischemic stroke patients who had antithrombotic therapy administered by end of hospital day two, with day one being the arrival day or date of arrival.

The exclusions for this measure are similar to the others, but this is a time sensitive measure. So it's probably a little bit more similar like our VTE prophylaxis measure that was discussed first today.

Excluded populations include patients less than 18 years of age, patients who have a duration of stay less than two days being calculated from the arrival date, patients who have length of stay greater than 120 days, patients with comfort measures only documented on the day of or day after arrival, patients enrolled in clinical trials, patients admitted for elective carotid intervention, patients who are discharged prior to the end of hospital day two.

And then a new data element for this measure, we exclude patients with IV or IA thrombolytic therapy administered at this hospital or within 24 hours prior to arrival at the hospital to account for possibly the drip and ship patients that may transfer in.

And then there is a reason exclusion patients with documented reason for not administering antithrombotic therapy by end of hospital day two. The evidence for this, when it was last endorsed in 2012, was high.

There are many clinical studies that demonstrate the benefit of early antithrombotic therapy in reducing stroke mortality and stroke related morbidity.

The recommendation, actually, which is Class 1 level of evidence A, recommends administration of an antithrombotic, preferably aspirin within 24 to 48 hours of stroke symptom onset.

MS. ANDERSON: And the eCQM mimics the chart-abstracted measure as well. And this measure has been in use the same as the previous measure that we talked about for the meaningful use program at HIQR.

CO-CHAIR TIRSCHWELL: And when you say has been in use, meaning it's sort of on their list of things, but there are no data available

to you?

MS. ANDERSON: Correct, correct.

CO-CHAIR TIRSCHWELL: Okay. So over to, I don't know, Jocelyn or Steve? Who's going?

MEMBER BAUTISTA: I'll present.

CO-CHAIR TIRSCHWELL: Great, thank you.

MEMBER BAUTISTA: All right, so going to evidence. The developer has not submitted any new evidence since the prior endorsement. And so the evidence back then was high, multiple randomized control trials have shown the benefit of aspirin in acute ischemic stroke. So I don't know that there's anything to discuss.

CO-CHAIR TIRSCHWELL: I would agree and that we don't need to vote either. We can move right on to gap, unless somebody has a comment or discussion? Why don't we go with gap?

MEMBER BAUTISTA: So you can see from the table that we just have the same issue. The mean hospital performance is 98 percent the last three years, since 2012. And the tenth

percentile as well stayed 95, 96 percent the last three years.

So just taking that at face value, there does not appear to be much of a gap. The whole issue of disparities comes up again. There's some, perhaps, evidence from the literature that there may be some disparities in performance on this metric, but the developer doesn't submit any specific disparity data.

CO-CHAIR TIRSCHWELL: Discussion about gaps? Seeing none, I think we should probably vote on this. The initial recommendation was for low gaps.

MS. OGUNGBEMI: We are now voting for a performance gap on Measure 0438. The options are high, moderate, low, and insufficient. Voting is open.

Voting is closed. The results are zero percent high, 13 percent moderate, 87 percent low, and zero percent insufficient. Measure 0438 does not pass on performance gap.

CO-CHAIR TIRSCHWELL: And so I think

we still need to go through the others because of the question about possibly going to reserve status versus non-endorsement. So let's proceed along to reliability.

MEMBER BAUTISTA: So for reliability, no new information was presented on reliability. There was previously presented, in 2012, reliability testing that showed high overall agreement rate, 97 percent, with only one data element less than 95 percent. So I don't know that we need --

CO-CHAIR TIRSCHWELL: So -- right. If there's no new data, do we have to vote? Or can we pass on the reliability vote also? Any discussion about this? And with there being no new data, I would say we just move on to the next, validity.

MEMBER BAUTISTA: There is new validity information presented, empirical validity testing on the measure score. They looked at over two million patient records, 1,300 hospitals, and found high, well positive

correlations with six other stroke measures indicating, I guess, what you would call convergent validity.

They do present some information on threats to validity. They present their exclusion data which seems reasonable. And the preliminary rating for validity was high.

CO-CHAIR TIRSCHWELL: Any discussion or comments on validity? Okay. Well then I would move to vote since there was some new data.

MS. OGUNGBEMI: We are now voting on the validity for Measure 0438. Options are high, moderate, low, and insufficient. Voting is open.

Voting is closed. The results are 57 percent high, 43 percent moderate, 0 percent low and 0 percent insufficient. Measure 0438 passes on validity.

CO-CHAIR TIRSCHWELL: Feasibility?

MEMBER BAUTISTA: So feasibility. This measure has been in use for many years. It is reliant on data abstraction, which does represent some burden. So the preliminary rating

for feasibility is moderate.

CO-CHAIR TIRSCHWELL: Discussion? Go ahead and vote. Oh, sorry. Yes, Steve.

MEMBER HUFF: I'll be quick. It's just -- I realize this measure's been in place a long time. It just seems like such an odd time measure. We have other measures where the granularity goes down to a minute, and this is hospital day two which, at least to this clinician, is not part of a standard timeframe.

It's hospital day two could be anywhere from what, 24 hours and one minute into admission up to one day, 23 hours, 59 minutes. And just an observation. It would seem to me there should be some consistency through the measures.

CO-CHAIR TIRSCHWELL: You know, I think that timeframe is based on some of the clinical trial data where that was the timeframe within which people were given antiplatelet agents and there was shown to be a clear benefit on outcomes. I don't know, developer? Oh,

sorry.

DR. SCHWAMM: Yes, so the CAST trial was the trial that demonstrated benefit if it was given within the first 48 hours. The original implementation of the measure, many years ago, required sites to put the precise time at which the first medication was given.

And the feedback at that time was it was too onerous. And so it was changed to second hospital day which very closely, if anything, biases the measure toward earlier, rather than later treatment and was felt by hospitals to be much less onerous. So that was the rationale for that change many years ago.

CO-CHAIR TIRSCHWELL: Any other comments or discussion? Let's go ahead and vote on feasibility then.

MS. OGUNGBEMI: We are now voting on feasibility for Measure 0438. Options are high, moderate, low, and insufficient. Voting is open.

Voting is closed, results are in. Twenty-two percent high, 78 percent moderate,

zero percent low and zero percent insufficient. Measure 0438 passes on feasibility.

CO-CHAIR TIRSCHWELL: Great. And then finally use and usability.

MEMBER BAUTISTA: So the metric is used in multiple programs for public reporting, accountability, for quality improvement. So it's clearly demonstrated usability.

CO-CHAIR TIRSCHWELL: Any discussion? Any unintended consequences that we wanted to review? Okay, let's go ahead and vote then.

MS. OGUNGBEMI: We are now voting on usability and use for Measure 0438. The options are high, moderate, low, and insufficient. Voting is open.

Voting is closed. The results are 83 percent high, 17 percent moderate, 0 percent low and 0 percent insufficient. Voting, sorry, Measure 0438 passes on usability and use.

CO-CHAIR TIRSCHWELL: Great. And now, so we now vote on overall endorsement. Which -- no.

MEMBER BAUTISTA: But it failed on gap.

CO-CHAIR TIRSCHWELL: It failed on gap, right?

MEMBER BAUTISTA: Yes. It failed on gap.

CO-CHAIR TIRSCHWELL: So we're voting instead on reserve versus non-endorsement? Okay. So not this vote. This is not the right vote. And could you just read out all the criteria for us real quick?

MS. OGUNGBEMI: Yes. We are now voting on endorsement maintenance, the potential for reserve status. If a measure is under endorsement maintenance review and did not meet importance to measure and report only due to lack of the performance gap, or the criteria 1B, does it meet criteria to create for potential reserve status?

High performance is likely due to actual improvement versus an issue with measure construction. There is strong direct evidence,

it's proximal to the desired outcome, there are high ratings for reliability and validity, possibly moderate, it demonstrates use as well as improvement. Voting is open, options are yes or no.

Voting is closed. Results are 91 percent yes, nine percent no. Measure 0438 does pass on the potential for reserve status.

CO-CHAIR TIRSCHWELL: Great. Let's move immediately to 2835, which is the companion eMeasure. And I think we could probably immediately skip over evidence.

DR. TERRY: But it's reserve status.

CO-CHAIR TIRSCHWELL: Oh, sorry. Never mind. Got it, right, because that one isn't fully endorsed, this one doesn't even get discussed. Okay. Good thing you guys are watching out for me.

All right, then. So then I guess we're then moving at a better pace to 0436, anti-coagulation therapy for Afib/flutter with its companion eMeasure 2833. Again, the joint

commission.

MS. KOLBUSZ: Again, basically the same on the chart based measure, stroke three, anti-coagulation therapy for atrial fibrillation flutter was last endorsed in 2012.

It's used in hospital and patient quality reporting, it's collected by the Paul Coverdell National Acute Stroke Registry and the Joint Commission certification programs.

It's used in meaningful use, all that still applies. This measure captures the proportion of ischemic stroke patients with atrial fibrillation flutter who are prescribed anti-coagulation therapy at hospital discharge.

The excluded populations are similar to the other discharge measures. Patients less than 18 years of age are excluded, length of stay greater than 120 days, comfort measures only documented, patients enrolled in a clinical trial related to stroke, patients admitted for elective carotid intervention, discharged to another hospital, patients who left against medical

advice or who expired, discharges to home for hospice care, discharges to home to another healthcare facility for hospice care.

And then we have a recent data element to exclude patients with a documented reason for not prescribing anti-coagulation therapy. This particular measure, since it's focusing on patients with a history or current finding of non-valvular atrial fibrillation, the recommendation is for anti-coagulation therapy at discharge, a little bit more potent than our antithrombotic measure. There's a large body of evidence to support that recommendation.

The original studies that we used to develop the measure were based on warfarin use. However, in recent years, novel oral anti-coagulant agents now referred to as the DOACs, the direct oral anti-coagulants, have been developed and approved by the U.S. FDA for stroke prevention and may be considered as an alternative to warfarin for select patients.

And those agents have all been added

to our list of acceptable drugs for inclusion in the numerator. That's it.

MS. ANDERSON: And Karen did mention that the eCQM is used for the meaningful use program and HIQR, same as the rest of our eCQMs for stroke. People have been attesting to it for meaningful use but we do not have sufficient data to actually do analysis of that.

CO-CHAIR TIRSCHWELL: Great, thank you. And so Ketan and Alex? Okay, great.

MEMBER BULSARA: So in terms of evidence, I mean, this is, it's just undergoing maintenance evaluation. It's a process measure. There's no new evidence.

I mean, anti-coagulation does reduce the risk of stroke well established. So not much to discuss there. I just thought something to think about for future sort of renditions of this to develop some sort of data regarding the timing consensus because there seems to be variability in that. But there's nothing to really discuss in terms of new evidence. So we can move --

In terms of gap, you know, this is one of those measures that is very important that I think that we should definitely maintain our endorsement of. But you know, there's not much of a gap in terms of further improvement.

So I think we may want to consider, unless Alex feels a little bit differently in his review, that this may be something that we may want to consider reserve status for.

CO-CHAIR TIRSCHWELL: Any further comments on gaps? With that then, no comments, let's go ahead and move to a vote.

MS. OGUNGBEMI: We are now voting on performance gap for Measure 0436. Options are high, moderate, low, and insufficient. Voting is open.

Voting is closed. Results are zero percent high, 26 percent moderate, 74 percent low, and zero percent insufficient. Measure 0436 fails on performance gap.

MEMBER BULSARA: In terms of reliability, as was pointed out earlier, the

numerator statement was ischemic stroke patients prescribed anti-coagulation therapy at discharge. Denominator was ischemic stroke patients with documented Afib and flutter.

And we've already gone over the exclusion criteria. Have no major issue in terms of reliability, but just out of interest, in terms of reliability testing, the numerator did change and it's been changed with the addition of new anti-coagulants. So I'm just curious as to how, what was the extent that that numerator changed?

CO-CHAIR TIRSCHWELL: So you're asking for data which we were not presented with?

MEMBER BULSARA: Just to get a sense for again, you know, so one of the criticisms in the past was not the new anti-coagulants weren't considered. Those are being considered now, but I think just out of more general interest, how did that change the numerator? And we may not have the data. And it's not pertinent to this discussion.

CO-CHAIR TIRSCHWELL: So you mean did performance improve with including, you know, how many people were getting the newer oral anti-coagulants initially when they weren't being counted for them?

MEMBER BULSARA: So my understanding is previously the patients that were being captured are the ones that were discharged on warfarin that now were added the patients with the new anti-coagulants. So how many more are we capturing?

CO-CHAIR TIRSCHWELL: Yes. When did the actual change in the numerator go into play?

DR. SCHWAMM: I think you're asking the question what was the gap in time between when the new agents have been approved for use and they show up on the form as an acceptable anti-coagulant?

CO-CHAIR TIRSCHWELL: Correct.

DR. SCHWAMM: And, you know, before Karen answers that, I'm just going to make the comment also that there are also patients who are

discharged on an anti-coagulant that's not approved for use for atrial fibrillation. Maybe they have a DVT and Afib.

They end up being excluded from the measure because it's an anti-coagulant so they're not appropriate for a different anti-coagulant. But they're not failing the measure, they just have another reason for anti-coagulation that puts them on an agent that isn't approved for use in AF.

CO-CHAIR TIRSCHWELL: But would the NOAC patients have been fails before the change?

DR. SCHWAMM: Before the change, they would have been a reason for exclusion, right, if they were on it. But it wasn't an available response option on the form. Am I saying that right?

MS. KOLBUSZ: It is accurate, what you said, Lee. Dabigatran was the first to be approved in 2010, it's a drug thrombotic inhibitor. And basically because the data element VTE prophylaxis was shared with other VTE

prophylaxis data measures in the core measures sets. For example, SCIP had two VTE measures.

No, I know. But basically because of that fact, in the data element VTE prophylaxis, we could not add a specific allowable value to capture dabigatran. So initially, what we did and we continue to do is we allow those patients to be captured in the reason data element, reason for no anti-coagulant, no VTE -- yes, so we allow it for the reason for anti-coagulation therapy at discharge. And they are captured, they are not excluded. They are in the numerator, I'm sorry about that.

Basically, they were added after 2010 and we've added them with each consecutive new drug that's been approved. We have Pradaxa, we have Savaysa, we have the edoxaban, we have apixaban, and that's all the ones that are currently FDA approved.

DR. SCHWAMM: So I think the point to make though is that once they get approved, at the very next -- once they've been approved, at

the very next update to the tool, they get added. So there is a couple of six months sometimes where a new agent's approved and it's not yet in the system, but it's a pretty short gap.

CO-CHAIR TIRSCHWELL: So it probably hasn't had a major impact on the performance data anyway?

DR. SCHWAMM: No.

CO-CHAIR TIRSCHWELL: Okay.

DR. SCHWAMM: The point I was trying to make before is, like, if you have for example these are only approved for non-valvular Afib. If you had valvular Afib and you were put on one of these agents, that's not really an approved indication.

So that wouldn't necessarily qualify. But as soon as they have a labeled indication for atrial fibrillation, they then go into the numerator.

CO-CHAIR TIRSCHWELL: Okay. Any other comments or discussion on reliability? Peter, no? So let's go ahead and vote.

MS. OGUNGBEMI: We are now voting on reliability for Measure 0436. The options are high, moderate, low, and insufficient. Voting is open. We're just waiting on one person. Oh, got it.

All right, the results are 13 percent high, 87 percent moderate, zero percent low, and zero percent insufficient. Measure 0436 passes on reliability.

MEMBER BULSARA: For validity, I think that's the area that I had the most trouble with this measure. The numbers, like Michael had pointed out on an earlier measure, the numbers didn't add up completely.

And I think the one aspect that didn't add up for me is there's a significant threat to validity with 50 percent of the patients being discharged to hospice. And the other issue was why discharge to another hospital was excluded.

CO-CHAIR TIRSCHWELL: So those were, yes, there was a math error in those data that the Joint Commission sent. And I think the new,

the updated, corrected data are available on the SharePoint site. And this is Measure 436.

So the old data, let's see, home for hospice was 52 percent, now it's one percent. So big difference there. I'm not exactly sure where they went sideways with the math, but they've fixed it.

I'm trying to see if there are any other ones for 436. I don't see any.

MEMBER BULSARA: I think with the new data, I mean, that would be more in line with our clinical practice. I mean 50 percent and one percent, that's pretty dramatic.

CO-CHAIR TIRSCHWELL: I think the discharge to another hospital probably suffers from the same issue at 35 percent.

MEMBER BULSARA: But it does bring up an issue, though. I mean, if we're going to present numbers, if we're going to make argument with numbers, I think we have to make sure that they're accurate so we don't spend needless time sort of re-reviewing things.

So I think it's imperative that the numbers accurately reflect what's actually going on. So it causes a lot of wasted time if they don't.

CO-CHAIR TIRSCHWELL: Yes. Any other comments or discussion on validity? Let's go ahead and vote then on validity.

MS. OGUNGBEMI: We are now voting on validity for measure 0436. The options are high, moderate, low, and insufficient. Voting is open.

Voting is closed and the results are 26 percent high, 74 percent moderate, zero percent low, and zero percent insufficient. Measure 0436 passes on validity.

CO-CHAIR TIRSCHWELL: Feasibility?

MEMBER BULSARA: No big issues on feasibility. There was a comment that sometimes data collection can be burdensome in terms of what's needed, but no issues from my perspective on feasibility.

CO-CHAIR TIRSCHWELL: Discussion? Let's go ahead and vote then on feasibility.

MS. OGUNGBEMI: We are now voting for feasibility on measure 0436. The options are high, moderate, low, and insufficient. Voting is open.

Voting is closed. Results are 35 percent high, 65 percent moderate, zero percent low, and zero percent insufficient. The measure passes on feasibility.

CO-CHAIR TIRSCHWELL: Usability?

MEMBER BULSARA: No issues on usability. There's seven credible organizations that are already incorporating this. So no issues on usability.

CO-CHAIR TIRSCHWELL: Discussion? Let's go ahead and vote on usability and use.

MS. OGUNGBEMI: We are now voting for usability and use for Measure 0436. The options are high, moderate, low, and insufficient. Voting is open.

Voting is closed. Results are 74 percent high, 26 percent moderate, zero percent low and zero percent insufficient. Measure 0436

passes on usability and use.

CO-CHAIR TIRSCHWELL: Okay, so that puts us, I think, in the same boat as last time where we failed on gap and so we can vote about reserve status. Getting good at this. Yes, Charlotte?

MEMBER JONES: If we as a committee move everything to reserve, we need to reconsider what reserve means in terms of effort, and time, and the fact that we could end up always turning every measure over to reserve status.

And I'm not saying which measures or not, but if everything that doesn't pass gap goes to reserve, then there's no reason to even vote. We should just set up an algorithm that if we identify no gap and we always vote to put it on reserve status, spare one vote.

CO-CHAIR TIRSCHWELL: Okay. Peter?

MEMBER SCHMIDT: So I disagree with that position. I don't think that it's up to us to keep the panel full of measures. But I do think that we should take the opportunity to

communicate to measure developers where we think the next challenge might be or where to go from when we've seen success with a measure, where they might want to take it.

CO-CHAIR TIRSCHWELL: Dave?

CO-CHAIR KNOWLTON: Yes, I concur with what Peter said. I think that we've got an unusual set of circumstances today with these. Part of what's keeping us in that because I certainly don't want us loading up reserve status things.

But I think we've got eMeasures that are parallel with these measures right now and I think it was a difficult call for us today. So I think that's why I was seeing more reserve status on these ones, at least from where I sit.

MEMBER FERZIGER: Not to just join a chorus, but I think there's another point to think about and that is, you know, one of the things we've been challenged all day today, and I hope NQF will think about this is the issue of scope.

It seems like we keep coming into these policy issues that aren't really, you know, at least what I expected to be the scope of our task here. And so, you know, really the sort of scarcity issue, where the resources are and so forth, that's a policy question.

We shouldn't just be thinking about the neurology measures, we should be thinking about maybe we shouldn't measure neurology at all because pediatric cardiology, you know, is so important.

So I think your point is really important, but I just don't think it's within our purview to, you know, address the policy issue and that if it seems reasonable because the world changes to keep these things in reserve and keep looking at them, then that would be our recommendation.

CO-CHAIR TIRSCHWELL: Okay, any other discussion before we go ahead and vote on reserve status for this measure versus non-endorsement?

MS. OGUNGBEMI: We are now voting on

the potential for reserve status on Measure 0436. The options are yes and no. Voting is open.

Voting is closed. Results are 96 percent yes, four percent no. Measure 0436 passes on its potential for reserve status.

CO-CHAIR TIRSCHWELL: Okay, thank you. And so that means no discussion for 2833 then. We have a break scheduled here on our agenda.

CO-CHAIR KNOWLTON: We already took it.

CO-CHAIR TIRSCHWELL: We already took that one? I thought that was a bonus break. Okay.

No, I just thought we added one in. I didn't think, I didn't know I was trading out my later break. Okay, well then we'll keep going. You're up, my partner.

CO-CHAIR KNOWLTON: We are going to start discussing 441, assessed for rehabilitation. Again, Joint Commission?

MS. KOLBUSZ: Okay, the chart, the base measure stroke and assess for

rehabilitation. It does have the eCQM companion that mirrors it. This measure captures the proportion of ischemic or hemorrhagic stroke patients assessed for who received rehabilitation services during the hospital stay.

The exclusions for this are very similar to the other measures. Patients less than 18 years of age, patients who have a length of stay greater than 120 days, patients with comfort measures only documented, patients enrolled in clinical trials related to stroke, patients admitted for elective carotid intervention, discharges to another hospital, discharges who left against medical advice, patients who expired, patients discharged to home for hospice care, patients discharged to healthcare facility for hospice care.

The basic rationale for the measure is that a large number of stroke patients who could benefit from rehabilitation services do not receive rehabilitation services. That estimate has been quite high in past studies.

And so the basic premise here is that at minimum, all patients' needs for rehabilitation services should be assessed some time during the hospital stay prior to discharge from the hospital.

CO-CHAIR KNOWLTON: So reviewers for this were Michelle and Ross.

MEMBER ZAFONTE: I think I'll start and Michelle will chime in and be helpful where I fall over the cliff. So the numerator is the ischemic or hemorrhagic stroke patients assessed or who received rehab services. The denominator is sort of all individuals with ischemic or hemorrhagic stroke.

Among the exclusion criteria that we were talking about we had some concerns with, and we understand why, because it's combined with other metrics.

It seems to me, just as a clinician, that you would want to know about the people at the longer length of stay when you're considering a rehab metric because that's the group of people

that would disproportionately benefit from it in preventing secondary comorbidities. But it's not within their derivation of the metric.

The evidence bases, this was reviewed in 2012. There doesn't seem to be that much in the way of new evidence for this. There is a series of clinical guidelines and smaller clinical trials, some outside the United States in Canada and Australia, others in other areas that suggest lower death and a lower rate of institutional care and some small studies that suggest improved functional status overall at one year.

So the preliminary rating of the evidence was moderate, but there hasn't been that much of a change since 2012.

(Off mic comments.)

MEMBER ZAFONTE: Do we need to vote, or --

CO-CHAIR KNOWLTON: There is no new evidence is there? No? Okay. No we don't.

MEMBER ZAFONTE: In the interest of

time and also sort of our Groundhog Day event, I'll go to the gap in care. Can we put that up again? Is that possible?

(Off microphone comments)

It's assessed or delivered.

MEMBER COTTER: How are people assessed for rehabilitation?

MEMBER ZAFONTE: So that comes into one of the definitional issues as we get down a little bit in this discussion. So one of the definitional issues is who contacts with them. And that has been expanded to include PAs and NPAs. And I suspect anybody contacting them from the rehabilitation team, there is a defined group of members who would define that assessment.

MEMBER KOENIG: I have a question first. Does a facility have to have a rehabilitation service? Is there any exclusion for, say, a small hospital does not have a rehabilitation service?

MEMBER ZAFONTE: I don't believe there is an exclusion there, but I turn to the metric

developers. I wasn't familiar that there was an exclusion.

MS. KOLBUSZ: If there was a small hospital that did not have a rehabilitation service, the physician, the attending physician would be considered a member of the rehabilitation team.

And of course, the expectation would then be that most likely, since there's no therapist there to do an assessment, that the attending would consider the rehabilitation needs of the patient before discharge.

MEMBER ZAFONTE: So as we turn to gap here, there's obviously been, you know, significant growth in the metric over the past five years, specifically from '13 to '14. But if we look at the 10th to 90th percentiles, you know, it's gotten pretty good.

It's up national aggregate rate of 0.97, 50th percentile is pretty much 100 percent, and the 10th percentile is at 89th percentile. So we're at the border, I think, of the same gap

issue again that we've been running into on a persistent basis.

CO-CHAIR KNOWLTON: Comments, questions? Michelle?

MEMBER CAMICIA: So there is some data published on disparities and rehabilitation --

CO-CHAIR KNOWLTON: Is your mic on? Speak into it.

MEMBER CAMICIA: There are some data available on disparities for stroke patients and post-acute levels of care, though that was not submitted. And again, we share the same disparities issue that we've discussed previously.

MEMBER ZAFONTE: Yes. And so that data's actually pretty compelling regarding long-term outcome, access to rehabilitation services. And so that might be the place that the gap was at. And I had asked just in our break, Karen, if they had any additional data for us.

And there may be some inter-hospital based differences, but I'm not sure about any

global differences that are available.

CO-CHAIR KNOWLTON: Anything from the developer on that? No?

MS. KOLBUSZ: That's correct, it was similar to the BT prophylaxis measure. Again, when we looked at the data that we did have for disparities in our database, that there was nothing at an aggregate level that was identified.

However, individual hospitals, we did note disparity that was significant for some. And the disparity would be about 2.1 percent of hospitals.

CO-CHAIR KNOWLTON: Okay. Alex?

MEMBER RAE-GRANT: The preliminary rating was low. Are you proposing a different rating for this than the preliminary?

MEMBER ZAFONTE: No. I think from my perspective, I'm probably proposing that it winds up in the so called emeritus status, again, because I think it's something we need to check and I would like to see more data on the

disparities.

And I understand why people have concerns with it, but I think that this disparities issue deserves to, at least, be put in that status and us take a closer look at it in the future.

MEMBER RAE-GRANT: So you would retain the low rating on the data?

MEMBER ZAFONTE: Yes.

CO-CHAIR KNOWLTON: Anybody else? Okay, let's vote on gap.

MS. OGUNGBEMI: We are now voting on performance gap for Measure 0441. The options are high, moderate, low, and insufficient. Voting is open.

Voting is closed. The results are zero percent high, nine percent moderate, 91 percent low, and zero percent insufficient. Measure 0441 fails on performance gap.

CO-CHAIR KNOWLTON: Once again, let's continue on.

MEMBER ZAFONTE: We're going to go

onto reliability. There has been an expansion in the definition regarding those individuals who are part of the rehabilitation team. Segue to our prior discussion a few minutes ago and the kinds of things that would be associated, i.e., a clarification of a clinical assessment.

There does not appear to be any major reliability issues, however this has been core crosswalked several times. And the bigger concern with this that has been pointed out several times again is that this is assessed at a data element level and not a measure score level. So it received a moderate reliability rating.

CO-CHAIR KNOWLTON: Questions or comments? Okay, let's vote.

MS. OGUNGBEMI: We are now voting on reliability for Measure 0441. Results are, or options are high, moderate, low, and insufficient. Voting is open.

Voting is closed. Results are nine percent high, 87 percent moderate, four percent low, and zero percent insufficient. Measure 0441

passes on reliability.

MEMBER ZAFONTE: We're going to go to validity next. The validity of this seems quite good. They did do a look with the seven other stroke metrics and tied it pretty well to the six core other stroke performance metrics supporting hypothesis that hospitals with high quality on one stroke metric tend to have relative high quality on others.

There are some threats to validity that I just wanted to go over again. I think Michelle and I's interest in this group of people who had the length of stay greater than 120 days, and then could we just confirm what the hospice rate was also on the re-done data on this?

(Off microphone comments)

MEMBER ZAFONTE: Yes, that's what I thought. Okay, because I have them both written down, one year, 50 percent which, of course, made no sense. Those are our only concerns with that.

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

MS. OGUNGBEMI: We are now voting on the validity for Measure 0441. The options are high, moderate, low, and insufficient. Voting is open.

Voting is closed. Results are 43 percent high, 57 percent moderate, zero percent low, and zero percent insufficient. Measure 0441 passes on validity.

CO-CHAIR KNOWLTON: Feasibility?

MEMBER ZAFONTE: Feasibility. There is some reported data collection burden for chart abstraction in this metric, but overall it was rated moderate and we happen to agree with that.

CO-CHAIR KNOWLTON: Questions or comments? Let's vote.

MS. OGUNGBEMI: We are now voting on the feasibility of Measure 0441. Voting is, or the options are high, moderate, low, insufficient. Voting is open.

Voting is closed. The results are nine percent high, 91 percent moderate, zero percent low and zero percent insufficient.

Measure 0441 passes on feasibility.

CO-CHAIR KNOWLTON: Usability?

MEMBER ZAFONTE: Usability of this is reasonably quite high. It's certainly been a metric that's been reported at a number of different organizations, I think over seven.

The results have been improved, with over 1,200 hospitals really obtaining a very high level of compliance with this. So the usability looks quite high.

CO-CHAIR KNOWLTON: Questions or comments? Vote?

MS. OGUNGBEMI: We are now voting on the usability and use of Measure 0441. Options are high, moderate, low, and insufficient. Voting is open.

Voting is closed. Results are 87 percent high, 13 percent moderate, zero percent low and zero percent insufficient. Measure 0441 passes on usability and use.

CO-CHAIR KNOWLTON: Comments on reserve status? Ross?

MEMBER ZAFONTE: So I think it's appropriate for reserve status for a number of different reasons, one related to the changing paradigm of care that we're all going under and this concern with utilization of post-acute services.

And here we have some at least with moderate evidence affecting long term outcome and re-hospitalization, as well as institutionalization for patients and as well as the disparities issues that Michelle raised that are pretty compelling in the literature, but we need more data about.

CO-CHAIR KNOWLTON: Anything further, Michelle, to that? Any other comments or thoughts? Peter?

MEMBER SCHMIDT: So just to comment to the measure developer, I know from my own research and my own activities that there are disparities in referrals. And there are, I just did a quick PubMed search, there is literature on this.

And if it had been included in the disparities section, I think you would find a lot of people very willing to -- say that there's a gap and perhaps that gap is addressing that 77 percent of hospitals who are not in the set who are currently reporting.

You know, there's definitely outcome benefit for referral, and I feel that there's a pretty -- you could very easily make the case that those disparities were there. And I wish you had, because it would have made a difference, I think, in the outcome today.

CO-CHAIR KNOWLTON: Other comments? Okay, we're voting on endorsement maintenance potential for reserve status.

MS. OGUNGBEMI: We are now voting for the potential for reserve status on Measure 0441. The options are yes or no. Voting is open.

Voting is closed, results are in. We have 96 percent yes, four percent no. The potential for reserve status of Measure 0441 passes.

CO-CHAIR KNOWLTON: We will be skipping the next measure because it is an eMeasure tied to the previous measure that did not meet gap. So we will be moving on to 2863, nimodipine treatment administered. Joint Commission again?

MS. KOLBUSZ: All right. The first measure that we are going to be discussing is C-stroke 6, which is comprehensive stroke 6, nimodipine treatment administered. And like the stroke measures that we've discussed all day, this is a new measure submission, the first time that it's coming to the committee for endorsement consideration.

This particular measure measures the proportion of subarachnoid hemorrhage patients age 18 years and older for whom nimodipine treatment was administered within 24 hours of arrival at this hospital.

It is an important measure to consider, the rationale for it being that nimodipine is the only calcium channel blocker

that's been found to be effective in the prevention of cerebral vasospasm which is a serious complication following subarachnoid hemorrhage that occurs in 30 percent to 70 percent of the patients and accounts for nearly 50 percent of the deaths surviving to treatment.

Therefore, we have included it in this particular set. The excluded populations include patients less than 18 years of age, patients who have a length of stay greater than 120 days, patients with comfort measures only documented on day of or after hospital arrival, patients enrolled in clinical trials related to stroke, and patients discharged within 24 hours of arrival at this hospital.

This measure was implemented effective January 1st, 2015. It was developed and is currently used by our disease specific care comprehensive stroke certification program which consists of 100 certified hospitals to date.

At the time that we submitted the measure, we had two quarters of data available

from actual implementation. Other data provided was from the pilot test of the measures. And I think that I'll just stop there and turn it over to the Chair.

CO-CHAIR KNOWLTON: Okay, thank you. Discussions will be Ron and Melody.

MEMBER RYAN: Okay, I'll take the lead and Ron will jump in when needed. Okay? So as far as evidence goes, the evidence is presented as a guideline from the American Heart Association, American Stroke Association with Level 1 Grade A recommendation for use of nimodipine for aneurysm subarachnoid hemorrhage.

There is also a Cochrane review, 16 studies which showed good likelihood of benefit. However, if the largest trial which had 906 patients in it is excluded, then the results are no longer statistically significant.

But that's probably just because of the numbers there. The algorithm points to high on level of evidence here. Ron, did you have anything else you wanted to add?

CO-CHAIR KNOWLTON: Questions or comments?

MEMBER KOENIG: I had a question.

CO-CHAIR KNOWLTON: Go ahead, Ron.

MEMBER KOENIG: For the developer, when you say subarachnoid hemorrhage, is it all hemorrhages or just secondary aneurysm?

MS. KOLBUSZ: The measure is intended for those patients that have an aneurysmal subarachnoid hemorrhage. That's provided in the rationale for the measure that we're focusing on those patients.

We do capture the initial patient population using the ICD 10 principal diagnosis code. There isn't a specific principal diagnosis code or any diagnosis code for that matter of non-aneurysmal subarachnoid hemorrhage.

Therefore, in order to really accommodate those patients, there needs to be some text documentation in the medical record that the patient was non-aneurysmal in order to address it in the measure which would be

addressed under the reason for not administering nimodipine treatment.

CO-CHAIR KNOWLTON: David?

CO-CHAIR TIRSCHWELL: So if the record says that this convexity subarachnoid hemorrhage was presumptively due to amyloid angiopathy, that would be enough to get you excluded from the measure?

MS. KOLBUSZ: We have in the reason data element for the measure, the reason for not administering nimodipine treatment, we did provide some stand alone reasons, non-aneurysmal is one. I believe the cerebral amyloid angiopathy was on there, I have to actually look it up and find it in my measures.

(Simultaneous speaking)

CO-CHAIR TIRSCHWELL: Well it specifically implies non-aneurysmal.

MS. KOLBUSZ: Okay, yes. Okay, so what I'm saying is though we would look for text documentation within the first 24 hours which is the timeframe for the measure because we can't

capture it through any other type of coded information such as an ICD 10 diagnosis code. So we would be looking for some other documentation.

CO-CHAIR TIRSCHWELL: Why does it have to be in the first 24 hours? Sometimes the workup can be prolonged, repeat angiograms.

DR. SCHWAMM: So I think the issue there is that the measure performance is was it provided within the first 24 hours.

CO-CHAIR TIRSCHWELL: Oh, okay. All right. That makes sense.

DR. SCHWAMM: So that's the issue with the notation provided. Yes, so the goal is to target aneurysm with subarachnoid hemorrhage. And there is an exclusion criteria for non-aneurysmal.

What Karen is saying is if there was an ICD-9 code from aneurysmal subarachnoid hemorrhage, they could to the exclusion based just on ICD-9 code because there is one they have to rely on text based description of why it's not appropriate in this case.

CO-CHAIR TIRSCHWELL: But we don't have to say in the chart I'm not treating this non-aneurysmal subarachnoid hemorrhage with nimodipine because it's totally inappropriate to do so?

DR. SCHWAMM: So you're asking do the magic words non-aneurysmal have to appear in the text? I don't know the answer to that.

MS. KOLBUSZ: Well, I think what you're asking is maybe twofold. A stand-alone reason we're looking for text documentation that the patient is non-aneurysmal.

We aren't requiring other linkage with nimodipine because we're considering it a stand-alone reason. So we're not looking for an explanation beyond non-aneurysmal. But there needs to be some documentation to identify those cases.

CO-CHAIR TIRSCHWELL: Okay, thank you.

CO-CHAIR KNOWLTON: Both of our discussants want to comment, so go ahead Melody.

MEMBER RYAN: So when you look at the

coding instructions that are given for it, the stand-alone reasons are non-aneurysmal subarachnoid hemorrhage, reversible cerebral vasoconstriction syndrome, and cerebral amyloid angiopathy. And all of the codes that are given are for non-traumatic also.

CO-CHAIR KNOWLTON: Ron, something further?

MEMBER KOENIG: Just for the Committee for review, one of the exclusions was discharge within 24 hours. And initially the information I think provided was about 60 percent. It seemed outrageously high. But when new data was presented by Christie today, it was sent out I think it comes down to one percent.

CO-CHAIR KNOWLTON: Anything further here? Questions, comments? Yes, sir?

MEMBER HUFF: The most common cause of subarachnoid hemorrhage in emergency departments is traumatic. It might just be easier to push if trauma had been entered as an exclusionary point.

MS. KOLBUSZ: The trauma cases would

not be. There are separate hemorrhagic codes for trauma and non-trauma. So they are excluded because they're not included in that group of initial hemorrhagic patients.

CO-CHAIR KNOWLTON: Ron, are you still up? Ketan?

MEMBER BULSARA: Just along what Steve's saying, do you really need to include other, I mean, I think it confuses matters if you, the more things you put into differentials because I think there's a higher likelihood you'll leave something out.

Why not put cavernomas in there too? Or put capillary telangiectasias and things to that extent? Why not just leave the exclusion criteria just to say aneurysmal subarachnoid hemorrhage?

(Off microphone comments)

MS. KOLBUSZ: The code is subarachnoid hemorrhage. So if you don't have a code or some other way to identify that it's aneurysmal or non-aneurysmal, the code isn't making the

difference either way. So we need another way to identify those cases.

MEMBER BULSARA: But there is a code for aneurysm and there is a code for brain aneurysm.

MS. KOLBUSZ: There is.

MEMBER BULSARA: So --

MS. KOLBUSZ: And it's unruptured. It's an unruptured code.

MEMBER BULSARA: But in order for the patients to qualify for this measure, you would have to have both ICD codes.

MS. KOLBUSZ: We have had feedback during the pilot test regarding the aneurysmal and non-aneurysmal patients as well as after from our comprehensive stroke centers.

And we've looked at a variety of ways to address their concern that because of the coding system, they're pulling in all the subarachnoid hemorrhages and how would we address these patients that are non-aneurysmal which is the largest group. I'm sure there's other

diagnoses that obviously nimodipine isn't appropriate for.

But the best way to do it in the measure construct and the way that the other measures are constructed and the way that the hospitals are familiar with abstracting was to provide in the abstraction guidelines for the reason data element guidance.

And we tried to make it easier and limit the burden of abstraction by making it a stand-alone reason so that they don't have to look for additional documentation.

CO-CHAIR KNOWLTON: Any further questions or comments? Ready for a vote. This is evidence.

MS. OGUNGBEMI: One moment. We are now voting on evidence for Measure 2863. The options are high, moderate, low, and insufficient. Voting is open.

Voting is closed. The results are 61 percent high, 39 percent moderate, 0 percent low and 0 percent insufficient. Measure 2863 passes

on evidence.

CO-CHAIR KNOWLTON: Gap?

MEMBER RYAN: So there's also a table similar to the other ones that has been presented. For this measure, there were two different kind of phases of testing.

The first was the pilot testing which went on from 2010 to 2013. It had 66 sites and 1,229 patients. And then when it was actually implemented in the first and second quarters of 2015, the first quarter there were 39 sites and 572 patients, and then 51 sites and 873 patients in the second quarter.

The averages were high for all of them, but when you look at that table, could you come up so everybody can see it, you do see increases from the pilot time to the time it was implemented.

And the range in the pilot testing went from 0 to 100. So that would indicate that there's a need there. You can see the percentiles are given for the actual

implementation phase. And still I think there's a pretty good need for that.

When we look at disparities, the information that's presented is that there's a difference in mortality and of patients based on race. Compared to white, Hispanics were least likely to have a disparity.

And then blacks and Asian and Pacific Islanders were most likely to have a disparity. There is also a disparity between women and men for aneurysm. But none of these data actually look at a disparity between who gets nimodipine and who doesn't get nimodipine.

So I think that part, the disparity part of the performance is, you know, is not there yet for this. But I do think there is a pretty big gap on performance. The preliminary rating on that I believe was moderate.

CO-CHAIR KNOWLTON: Questions or comments? Let's vote.

MS. OGUNGBEMI: We are now voting for performance gap on Measure 2863. Options are

high, moderate, low, insufficient. Voting is open.

Voting is closed. Results are 13 percent high, 87 percent moderate, 0 percent low, and 0 percent insufficient. Measure 2863 passes on performance gap.

CO-CHAIR KNOWLTON: Reliability?

MEMBER RYAN: Okay. To briefly restate, the numerator is all patients for whom nimodipine treatment was administered within 24 hours of arrival. The denominator is all subarachnoid hemorrhage patients.

Exclusions from the denominator are patients less than 18 years of age, patients who have a length of stay greater than 120 days, patients with comfort measures only, patients enrolled in clinical trials, patients discharged within 24 hours of arrival at the hospital, and those are the exclusions.

So as I said, looking in the coding manual you can, it actually shows you that the measures, the codes that are given are for non-

traumatic subarachnoid hemorrhage, and also those other ones are stand-alone reasons for not administering nimodipine treatment. So non-aneurysmal, reversible vasoconstriction, and cerebral amyloid angiopathy.

So that was that. The testing was tested only at the data element level which will have implications for our voting. The inter-rated reliability testing was done at 12 sites with 281 records. The agreement was very high at 95 percent for everything except admitting time which was 82 percent. So all of that was actually high.

There were a number of hospitals but only 281 records that we used, so that does seem a little bit low. The preliminary ranking on that is moderate.

CO-CHAIR KNOWLTON: Questions or comments? Yes, Jim?

MEMBER BURKE: So I think up until this point we've mostly been talking about ischemic stroke in our ICDE codes which are

usually pretty well validated. Is there data particularly on the new, because it looks like under ICD-10 the subarachnoid codes change compared to before. Is there any data on the reliability of the codes as opposed to the data elements themselves?

MEMBER RYAN: So probably this is best discussed by the developers. But they do provide under threats to validity I think some information about crosswalking between ICD-9 and ICD-10. And I don't know what the actual name of that document is that is more involved, the one that comes after the brief summary statement.

CO-CHAIR KNOWLTON: Other comments or questions? Let's vote.

MS. OGUNGBEMI: We are now voting on reliability for Measure 2863. Options are high, moderate, low, and insufficient. Voting is open.

Voting is closed. Results are 9 percent high, 91 percent moderate, 0 percent low and 0 percent insufficient. Measure 2863 passes on reliability.

CO-CHAIR KNOWLTON: Validity?

MEMBER RYAN: Okay, for validity this is at the measure score level. And they were tested empirically. The first hypothesis was that there would be two measures that would be highly aligned on hemorrhagic stroke.

So the one is this one, 06 nimodipine treatment administered. The other one is 03 which is severity measurement performed for subarachnoid hemorrhage and intercerebral hemorrhage.

So looking at the severity in those two you would think would be highly correlated. They weren't. I don't know why, they just weren't.

Hypothesis two was that hospitals who do well on one stroke measure are likely to do well on the others and that one was confirmed with high correlations between those.

One concern we had on the workgroup phone call was that there was a high percentage of people that seemed to be discharged within 24

hours as previously mentioned by Ron. That actually is quite low, it's 1.42 and that was a mathematical issue that was corrected this morning.

Meaningful differences, there is a plan for meaningful differences. The range again, as we said, was pretty wide. So it seems like there's a possibility for good, meaningful difference there.

For missing data there is a plan, but we don't know how often data are actually missing for each one of those elements. And so the preliminary ranking here was moderate.

Some of the reason for that ranking though was because of the concern about so many people being discharged within 24 hours which we know is erroneous.

CO-CHAIR KNOWLTON: Questions or comments? Let's vote.

MS. OGUNGBEMI: We are now voting on validity for Measure 2863. Options are high, moderate, low, and insufficient. Voting is open.

Voting is closed. Results are 4 percent high, 91 percent moderate, 4 percent low, and 0 percent insufficient. Measure 2863 passes on validity.

CO-CHAIR KNOWLTON: Feasibility?

MEMBER RYAN: The source of data here is either electronic or paper medical records. And it does involve manual chart review which would mean having to go through the charts, you know, manually. And so that could be burdensome.

The developer estimated that the time for the measure set was 45 minutes per record. Surprisingly I think that's a bargain at only $3.50. But it does seem like these are data that are all available generally in the medical record and could be extracted.

CO-CHAIR KNOWLTON: Questions or comments? Let's vote.

MS. OGUNGBEMI: We are now voting for feasibility of Measure 2863. Options are high, moderate, low, and insufficient. Voting is open.

Voting is closed. Results are 13

percent high, 83 percent moderate, 4 percent low and 0 percent insufficient. Measure 2863 passes on feasibility.

CO-CHAIR KNOWLTON: Usability.

MEMBER RYAN: So the measure is currently being used for care certification for comprehensive stroke centers. In the future it is planned for public reporting and external benchmarking, but there's not a lot of detail or timeframe on that.

Whether or not that's desirable it's possible. I think the public might get more out of knowing that somebody is, you know, center stroke certified rather than specifically about this issue. So that was preliminarily ranked moderate.

CO-CHAIR KNOWLTON: Questions or comments? Let's vote.

MS. OGUNGBEMI: We are now voting on the usability and use of Measure 2863. Options are high, moderate, low, and insufficient. Voting is open.

Results are 13 percent high, 83 percent moderate, 4 percent low, and 0 percent insufficient. Measure 2863 passes on usability.

CO-CHAIR KNOWLTON: Overall? Well, we're now going to vote on the overall acceptability of the measure. Suitability for endorsement.

MS. OGUNGBEMI: Yes, we are now voting for the overall suitability for endorsement of Measure 2863. The options are yes or no. Voting is open.

Voting is closed. Responses are 96 percent yes, 4 percent no. Measure 2863 passes at suitability for endorsement.

CO-CHAIR KNOWLTON: I'm going to talk a bit with David and what we're going to do is we're going to defer the next issue to tomorrow because we want to have our robust discussion and then we've about zoned everybody out I think.

So I'm going to ask the operator to open the phone line for anybody on the phone that would like to comment, and I will also open it to

our guests if anybody would like to comment about issues that have been discussed today. Operator, would you open the phone line please?

OPERATOR: At this time if you would like to make a comment, please press star and then the number one. No comments from the phone line.

CO-CHAIR KNOWLTON: Anybody here? Nobody here either. Okay. So thank you for an assertive first day. We will be, those of you who are joining us, we're dining at Georgia Brown's at 6:15 I believe.

(Off microphone comments)

CO-CHAIR KNOWLTON: 6:15, correct? And we will be reconvening here tomorrow for breakfast at 8:00 a.m. Be here bright eyed and bushy tailed, we've got a lot to do tomorrow. Thank you.

(Whereupon, the meeting in the above-entitled matter was concluded at 5:03 p.m.)

D

H

J

W

C E R T I F I C A T E

This is to certify that the foregoing transcript

In the matter of: Neurology Standing Committee

Before: NQF

Date: 04-04-16

Place: Washington, DC

was duly recorded and accurately transcribed under my direction; further, that said transcript is a true and accurate record of the proceedings.

Neal R Gross

Court Reporter