NATIONAL QUALITY FORUM
+ + + + +
NEUROLOGY ENDORSEMENT MAINTENANCE
STEERING COMMITTEE
+ + + + +
WEDNESDAY
JUNE 20, 2012

+ + + + +

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

PRESENT:
DAVID KNOWLTON, MA, Co-Chair, New Jersey Health Care Quality Institute
DAVID TIRSCHWELL, MD, MSc, Co-Chair, University of Washington
A.M. BARRETT, MD, Stroke Rehabilitation Research, Kessler Foundation

WILLIAM BARSAN, MD, University of Michigan Health System
JOCELYN BAUTISTA, MD, MBA, Cleveland Clinic
RAMON BAUTISTA, MD, MBA, University of Florida HSC/Jacksonville
GWENDOLEN BUHR, MD, Duke University
GAIL AUSTIN COONEY, MD, FAAHPM, Hospice of

Palm Beach County/Spectrum Health Inc.
JORDAN EISENSTOCK, MD, CPE, UMass Memorial Medical Center
RISHA GIDWANI, DrPH, Stanford University Medical Center
DAVID HACKNEY, MD, Beth Israel Deaconess Medical Center

GREGORY KAPINOS, MD, MS, North Shore-LIJ Health System

PRESENT(Cont'd):
MICHAEL KAPLITT, MD, PhD, Weill Cornell Medical College
DANIEL LABOVITZ, Montefiore Medical Center
THERESE RICHMOND, PhD, CRNP, FAAN, University of Pennsylvania, School of Medicine
JACK SCARIANO, MD, PLLC, Neurologist, Fort Sanders Parkwest Medical Center and Tennova West Medical Center
RAJ SHETH, MD, Nemours Children's Clinic
JOLYNN SUKO, MPH, Virginia Mason Medical Center
JANE SULLIVAN, PT, DHS, MS, Northwestern University Feinberg School of Medicine
FREDRIK TOLIN, MD, MBA, FACS, Humana
MARY VAN de KAMP, CCC-SLP, RehabCare, Kindred Healthcare
SALINA WADDY, MD, National Institutes of Health

NQF STAFF:
HELEN BURSTIN, MD, Senior Vice President, Performance Measures
ANN HAMMERSMITH, JD, General Counsel
KAREN JOHNSON, MS, Senior Director, Performance Measures
SUZANNE THEBERGE, MPH, Project Manager,

Performance Measures
JESSICA WEBER, MPH, Project Analyst, Performance Measures
REVA WINKLER, MD, MPH, Senior Director, Performance Measures

ALSO PRESENT:

MARK ANTMAN, DDS, MBA, American Medical Association

JOSEPH DROZDA, JR., MD, AMA-PCPI*

KENDRA HANLEY, MS, AMA-PCPI

DIEDRA JOSEPH, AMA-PCPI

IRENE KATZAN, MD, MS, Cleveland Clinic

KAREN KOLBUSZ, The Joint Commission

LEE SCHWAMM, Partners HealthCare System, Inc.

DAVID SEIDENWURM, MD, AMA-PCPI

MARK STEWART, American Heart Association

SARAH TONN, American Academy of Neurology

ANN WATT, MBA, The Joint Commission

LAURA YODICE, MPH, MHA, American Medical Association

*present by teleconference

CONTENTS

CONTENTS(Cont'd)

CONTENTS(Cont'd)

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

9:39 a.m.

OPERATOR: Welcome to the conference. Welcome to the Neurology Steering Committee meeting.

Please note today's call is being recorded.

Please stand by.

MS. JOHNSON: Okay. Good morning, everybody, and welcome to the Neurology Steering Committee in-person meeting.

So, I just wanted to say thank you very much for all the work that you have done in looking at our measures so far and for flying in to beautiful, warm, balmy Washington, D.C.

I am Karen Johnson. I am the Senior Director for this project. And I think I have sent you a few emails along the way. So, it is nice to meet you guys.

What we are going to do this morning first thing is just go through and have brief introductions from everybody. I am probably

supposed to introduce our two Co-Chairs Dave and David, but I am going to let them do that. We will just go around and do just a very quick introduction.

And then, I think Ann will also help us with the COI thing.

MS. HAMMERSMITH: We combine introductions and COI. So, we will do it that way.

CO-CHAIR TIRSCHWELL: Okay. So, my name is David Tirschwell. I am a stroke neurologist. I work at Harborview Medical Center, which is part of the University of Washington in Seattle, Washington.

As a neurologist, a clinician, I work with mostly stroke patients, but I am very deeply committed to quality measures. I really appreciate all of you participating and donating your time to this process as well.

Should I say something about disclosures, too, now? No? Okay.

CO-CHAIR KNOWLTON: I am Dave Knowlton.

I am the other Co-Chair. I am the President and CEO of the New Jersey Health Care Quality Institute.

I also served on the first with some other folks here, Mary and some other people, on the first Stroke Committee that took on these measures.

The other thing that especially qualifies me for this is I am a stroke survivor. So, I represent purchaser and consumer interests.

We say nothing? I don't have anything to disclose. So, I would just say that.

MS. HAMMERSMITH: We will get there.

CO-CHAIR KNOWLTON: Okay. So, Ann?

MS. HAMMERSMITH: Good morning, everyone.

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

What we are going to do now is go through disclosures of interest. We will have you introduce yourselves at the same time.

If you recall some time ago, you filled out a form for us where we asked you a number of questions. We have reviewed those forms as part of the nomination and seating process on the Committee.

For those who are on the Committee, what we like to do, once you are on the Committee, in the first public meeting we like you to disclose anything that you think is relevant to your service before the Committee. Just because you disclose does not mean that you have a conflict. It is simply a disclosure.

I want to remind you of just a few basic principles. You sit as individuals on this Committee. You are individual subject matter experts. That is why we chose you to serve.

Sometimes Committee members in many of our Steering Committees, in perfectly good faith and innocently, will say, "I am John Doe and I am here representing the American Association of...." -- fill in the blank. Actually, you are not here representing

anybody but yourself. We are interested in your expertise.

The other thing that I want to remind you about is also related to things that I have heard Committee members say, again, innocently and in good faith. They will say, "I have no financial conflict of interest." We are, of course, interested in financial conflicts of interest and financial disclosures, if they are relevant.

But in this world that we work in, many of you serve as experts on committees. You are active in related activities to what we are talking about before the Committee, where there is no money involved at all. You do it as a volunteer, as you do for us. So, to the extent you have had any service like that, that is relevant. We ask you to disclose that.

We are particularly interested in research funding, grant funding, and any consulting that you have done, if it is

relevant to what is before the Committee.

So, what I am going to do is ask you to go around the table. I am going to start with the Co-Chairs. Introduce yourself. Tell us who you are with and if you have anything to disclose.

CO-CHAIR TIRSCHWELL: David Tirschwell again. I participate in research that is funded by both industry and NIH, although I don't think it is related. I do serve on a task force for the American Stroke Association, but that is not involved with any of the development for quality measures, though I am a user of many of the stroke quality measures.

CO-CHAIR KNOWLTON: I am Dave Knowlton, as I told you, and I don't have any conflicts.

MEMBER BUHR: My name is Gwendolyn Buhr. I usually go by Gwen. And I work at Duke University. I am a geriatrician. I spend most of my time working in long-term care and teaching various learners at Duke. And I

don't have any conflicts.

MEMBER COONEY: I am Gail Cooney. I work in hospice and palliative medicine in West Palm Beach, Florida. I am, actually, a reformed neurologist and have boards in pain and hospice and palliative medicine, which is now a subspecialty of neurology. And I don't believe I have any conflicts of interest.

MEMBER VAN DE KAMP: I am Mary van de Kamp. I am Senior Vice President of Clinical Operations for Rehab Care and Kindred. And my role is to oversee all the clinical education within long-term care, hospitals, births, and skilled nursing. I am a speech and language pathologist as well, probably most importantly. And I have nothing to disclose.

MEMBER TOLIN: Fred Tolin. I am Vice President with Humanity Insurance Company. I am involved with medical policy development and implementation. Other than working for an insurance company, I have nothing else to disclose.

(Laughter.)

MEMBER LABOVITZ: I am Daniel Labovitz. I am a stroke neurologist at Montefiore Medical Center for the Albert Einstein College of Medicine. I am a card-carrying member of the American Heart Association and use the quality measures, serve on a task force.

MEMBER SHETH: Raj Sheth, Nemours Clinic, affiliated with Mayo Clinic, Jacksonville. I am a pediatric neurologist, do mainly epilepsy. In that context, I have served both in industry and in the American Academy of Neurology and the American Epilepsy Society with task forces and other committees such as that.

MEMBER R. BAUTISTA: Hello, everyone.

Ramon Bautista from the University of Florida in Jacksonville. I am an Associate Chair and Associate Professor of Neurology and head the epilepsy program over there.

I was on the Board of Directors of Shands Hospital and the University of Florida

in Jacksonville. Aside from that, no other conflicts to disclose.

MEMBER SULLIVAN: Good morning.

I am on the faculty of the Feinberg School of Medicine at Northwestern University in Chicago. I am a physical therapist.

I have industry funding and NIDRR funding for stroke-related rehabilitation research, and I serve the APTA, American Physical Therapy Association, on their outcomes as a consultant.

MEMBER SUKO: Good morning.

I am Jolynn Suko, and I serve as the Administrator for the Neuroscience Institute at Virginia Mason Medical Center, which is a regional medical center in Seattle, Washington. My background is quality and quality measurement and administration. And I have nothing to disclose.

MEMBER RICHMOND: Hi. I am Teri Richmond. I am a Professor of Nursing at the University of Pennsylvania School of Nursing.

I am also a Professor of Nursing and Surgery at the Division of Trauma at the Furman School of Medicine. I have nothing relevant to disclose.

MEMBER GIDWANI: Good morning.

I am Risha Gidwani from Stanford University Medical Center. I am a health services researcher within neurosciences over there.

I have had past scientific consulting relationships with a variety of large multinational pharmaceutical companies, but none related to neurology. So, I believe there is no conflict of interest.

MEMBER BARRETT: Good morning.

I am Anna Barrett. You can also call me "A.M." if you like. I am from Kessler Foundation in New Jersey, which has been involved in the development of some neuro qual measures. I am the Director of Stroke Research there, and I also play a role in the Kessler Institute of Rehab, helping them to

form outcomes and assess outcomes.

And my conflicts of interest are, I think, related to research funding from the NIH, Kessler Foundation. The Kessler Foundation has received research funding for me from a company called O'Brien Technologies, which is now defunct. And four years ago, we had an investigator-initiated grant, I think, from Pfizer.

MEMBER KAPINOS: Good morning.

I am Greg Kapinos. I am a neuro intensivist at the North Shore-NIJ in New York. Nothing to disclose.

MEMBER BARSAN: I am Bill Barsan. I am an emergency physician at the University of Michigan. I have been involved in stroke research since the 1980s and am an active member of the stroke team at the University of Michigan.

As far as disclosures, I am the principal investigator for the NINDS-funded Neurologic Emergency Treatment Trials Network

and have another cooperative award from NIH and FDA to look at adaptative clinical trial designs.

MEMBER WADDY: I am Salina Waddy. I am a stroke neurologist and neurogeneticist. I am at the National Institutes of Health as a Program Director for Health Disparities within the Office of Clinical Research. I am on numerous trans-NIH as well as trans-agency federal groups. I don't know if you need to know the names of those individually, but I don't think I have anything else to disclose.

MEMBER EISENSTOCK: Hi, everyone.

I am Jordan Eisenstock. I am from UMass Medical Center in Worcester, Massachusetts. I am double-boarded in both neurology and psychiatry. Among a few different administrative titles, I am also the Director of Neuro Rehabilitation for UMass's System Rehab Hospital. I don't have any disclosures.

MEMBER J. BAUTISTA: I am Jocelyn Bautista. I am an epilepsy neurologist at the

Cleveland Clinic. I also serve as the Quality Officer for our Neurological Institute.

I have participated, along with the American Academy of Neurology and the American Epilepsy Society, in creating evidence-based guidelines for epilepsy, but nothing directly related to the measures today.

MEMBER SCARIANO: Yes, hi. I am Jack Scariano. I am a Board-certified neurologist from Knoxville, Tennessee. I have been in practice now, let's see, 35 years, and I have had a whole lot of experience with both urban and also rural patients. And I have no conflicts.

MEMBER KAPLITT: Hi. I am Mike Kaplitt. I am the Vice Chair for Research and Residency Director and Director of Functional Neurosurgery in the Department of Neurological Surgery at Weill Cornell in our Medical College in New York.

I have NIH and Defense Department funding for my basic research lab, as well as

funding from a foundation called the JPB Foundation for Parkinson's research. I have been involved in a variety of industry-sponsored clinical trials, but not in the past year. And I have no other real conflicts. I spent seven years on the NINDS Clinical Trial Section, but, again, that ended over a year ago. That's it.

MEMBER HACKNEY: I am David Hackney. I am a neuroradiologist at Beth Israel Deaconess Hospital in Boston and at Harvard Med School.

As a clinical neuroradiologist, I guess I make money every time somebody does an imaging study for a patient with strokes. So, that is probably my primary conflict.

(Laughter.)

I have administrative responsibilities in the Radiology Department. I have research funded by NIH and DoD. I am on far too many committees for radiology societies, all of which in one way or another are interested in imaging of neurologic disease and stroke. I

don't have any industry funding or industry-paid consulting. I am a current and past member of Study Sections for NIH and have done Study Section work for DoD. And I have done guidelines committee work for the American Society of Neuroradiology and the American College of Radiology.

MS. HAMMERSMITH: Thank you for making those disclosures.

Are there any Committee members on the phone?

(No response.)

No? Okay.

Do you have any questions of me about conflicts of interest? Or is there anything that you want to discuss with each other, based on the disclosures this morning?

(No response.)

Okay. Thank you. Have a good meeting.

MS. JOHNSON: Thank you, guys.

And let's go ahead and have the NQF staff introduce themselves.

MS. THEBERGE: Good morning, everyone.

I am Suzanne Theberge. I am the Project Manager on this project. I think I have spoken with all of you on the phone, by email, and in person. It is great to meet you all.

And I just wanted to add that we have flash drives available with all the meeting materials, if anybody is having trouble getting online or needs them on a flash drive.

Jessica?

MS. WEBER: I am Jessica Weber. I am a Project Analyst at NQF.

DR. WINKLER: Good morning.

I am Reva Winkler. I am a Senior Director of Performance Measures at NQF.

DR. BURSTIN: Good morning, everybody.

I am Helen Burstin. I am the Senior Vice President for Performance Measures at NQF. I get to oversee this lovely process across all disciplines and cross-cutting areas.

So, thank you for your service. We know

there is a lot of work involved in reviewing these measures, and I just want to thank you for your time and your effort.

We will be here to provide support to the team, if there are any questions that come up, and there are always questions that come up. As measures get more and more complex, these decisions also get more and more complex.

So, thanks, everybody.

MS. JOHNSON: Thank you.

So, let's go ahead and start with just some real basic notes about the process that we are going to go through today. Or, actually, we are going to start with tools.

So, even though Dave has been teasing us about the trees that we have cut down for the materials that we have provided, we really have three hard-copy sets of things that we gave you.

One is the meeting agenda. So, that will tell you the order in which we will be

proceeding throughout the meeting.

And then, the second thing is the summary document. It is about a 62-page document that briefly summarizes each measure, gives the results of those preliminary evaluations that each of you did, and then, also, includes the notes and summaries that we took from our Work Group meetings. So, hopefully, that will be useful to you.

Risha is looking at me with a question.

MEMBER GIDWANI: Do you have any extra copies of that document?

MS. JOHNSON: Yes, we do. So, if anybody else needs any of these documents, just raise your hand and Jessica can help you.

The third thing that I want to go over very quickly is our Quick Guide. It is the NQF Evaluation Quick Guide. It is a little, four-page document. We are not going to read it all, but really this guide, we wanted to give you an hard copy that you could keep in front of you just to remind you of the

different criteria and all the different things that you will be thinking of as you go through these measures.

So, again, there are four main criteria that we will be looking at for each of the measures. Importance to measure and report is the first criteria. It is a must-pass. Underneath that overarching criterion is three subcriteria, and each of those three also are must-pass criteria. And those are high impact, evidence, and performance gap. And notice that we did switch evidence and performance gap around a little bit, and that is how we will go through today in our voting and such.

A reminder about evidence: depending on whether the measure is an outcome measure or a process or structure measure, it will change how you think about evidence and the kinds of things that you would expect to see.

I don't believe we have any outcome measures on our agenda for today. So, for all

of the measures today, you will be considering the quantity, quality, and consistency of evidence. Okay?

On the second page, well, I think the second and the top of the -- all of the second page is just what I told you. And again, I am going through these very quickly, but there will be time for questions, if you have any.

The second major criteria is scientific acceptability of measure properties. Again, that is a must-pass criterion. Underneath that criterion is two subcriteria that are also must-pass. Those are reliability and then validity.

So, under reliability, that is where you think about things like the specifications and, also, thinking about how they tested measures for reliability. Was the method appropriate? Were the results adequate to make you believe that the measure is reliable?

For validity, you are thinking about specifications there as well, but you are

thinking about it in terms of how the specs link to the evidence. So, you want to think about do the specifications actually reflect the evidence in the literature. You will be thinking about the validity testing, including the method and the scope, as well as the results. But you will also be thinking about things like exclusions, if risk adjustment was done, was that done well and was it adequate? You will be thinking about meaningful differences as well as comparability of data sources, if that is applicable.

We will be asking you to vote on reliability and validity separately. And then, we will use decision logic to see if a measure passes the scientific acceptability criterion.

The third and fourth criterion are listed on the last page of the guide. Usability, which reflects whether or not you believe that a measure is meaningful, understandable, and useful for public

reporting and accountability, as well as for quality improvement efforts.

And then, feasibility really gets to how usable, or not usable, but how the burden behind collecting data for a measure is, and is it really implementable? So, that is what you are thinking about in terms of feasibility.

And then, finally, there will be votes for overall suitability for endorsement. And I will walk you through those steps as well in a few minutes.

No. 5 on this sheet is comparison to related or competing measures. That is something that we will do after we go through the individual measures. So, we will look and evaluate each individual measure and then, if necessary, think about whether there are other measures out there that are either pretty much head-to-head competing with those measures, in which case you would have to consider if there is a best-in-class measure, or if there are

related measures, you would want to be thinking about how different are those measures. You know, do they have the same list of ICD-9 codes to describe a stroke, for example, those kinds of things. So, you are thinking about the details of the specs when you are thinking about harmonization of related measures.

So, let me stop right there and see if there are any questions on this very, very brief run-through of the evaluation criteria.

MEMBER KAPINOS: I'm sorry.

MS. JOHNSON: Yes?

MEMBER KAPINOS: I am surprised; you just said today, of all the measures, there are no health outcomes. I was in Work Group 4. So, I didn't look at all of them, but, briefly, before coming here, I saw that there was one that was measuring mortality. To me, that it a health outcome.

MS. JOHNSON: Right. That will be tomorrow.

MEMBER KAPINOS: Oh, that's tomorrow? Okay.

MS. JOHNSON: Right, right.

So, what we tried to do is -- we have a very packed schedule today -- what we tried to do was put all the process measures together and, also, all the measures from The Joint Commission from the AHA-ASA and AMA-PCPI, hopefully, to try to help them out, so that they only have to come one day. We may not get through everything. So, we may have to ask them to attend again tomorrow. But we will see how that goes.

And then, we will leave our brain space for tomorrow for those outcome measures that in some ways are a little bit more challenging, just because of the risk-adjustment issues that come up on those.

There are also measure submission materials. You guys have seen all of those. We didn't print them out for you. They are available either on the SharePoint site or on

the flash drives that Suzanne and Jessica have been passing around.

And then, finally, later on today, I am hoping to get you some comparison tables that will help as we go through the related/competing discussions that we will have at some point during the meeting.

Okay. Could we go to the next slide, please? So, I have already hinted at this, but, basically, the process is discuss, then vote. So, for each measure, we will be having eight votes. Okay? So, we will vote on all of the must-pass criteria.

So, what we will do is we will begin and talk about and discuss impact. And then, we will ask you to vote on whether or not you believe the measure has impact. Okay.

And then, we will go to evidence, discuss it, and then vote on it. We will work our way through.

If a measure fails a must-pass subcriterion, we will stop at that point. So,

if something fails at evidence, we will not go on to discuss opportunity for improvement or validity, reliability, that sort of thing. Okay?

For the first measure, what we will do is go through a little more slowly than we might on the other measures. As necessary, I will have slides up and walk you through some of the criteria. So, we are trying this, instead of giving you a 15-to-20-minute kind of encapsulation of that orientation tutorial that we did.

So, we will see how this works. If you think it is not working, let us know and we will change it for the next time around.

We will have, roughly, 15 minutes per measure. That is not a lot of time. So, we will do our best as we go through the day.

You should already have in your hands clickers. You will use those clickers to vote. When we get ready to cast our first vote, we will describe that a little bit more.

So, let's go to the next slide. A few housekeeping details. You have already managed to use your microphones. So, that is great.

This meeting is open to the public and is being recorded and transcribed as well. It is pretty important to have only one microphone on at a time. So, if you can possibly remember when you are done speaking, turn your microphone off, so that someone else's microphone will come on.

Your name tents in front of you are your signal that you would like to speak. So, if you will take your name tag when you have a question or a comment and just put it on its end, that is the flag to the Co-Chairs that you would like to speak. Yes, exactly. Exactly.

Our Co-Chairs are going to facilitate the meeting. So, NQF staff will be here to help if we need to, but it will mostly be facilitated by the Co-Chairs and you guys as

you go through the process. So, we are here to help, but you guys are the ones doing the evaluations.

We do, obviously, plan to have lunch and some breaks throughout the day. We are going to try to stick to those timeframes that we have for those. We may have to, depending on how fast we are going through, we may have to say, instead of a 15-minute break, let's take a 10-minute, something like that. So, we will see how the day goes and see if we need to swap out or not.

Restrooms, if you haven't already found them, are out in the hall. I am told that the men's room is very close to the women's room. So, out in the hall.

And finally, internet access, "Guest" and "NQF Guest" is the login and password.

Let's stop one more time and see if we have any questions. If not, we will start with our first measure.

Suzanne has reminded me -- I am sorry to

the developers -- we have asked our developers, and we have three different groups that are here today, we have asked each of the developers to spend two to three minutes each summarizing their measures that we will be thinking about today and tomorrow.

So, we don't have a list. How about AMA-PCPI?

MS. JOSEPH: Good morning, and thank you for the opportunity to introduce the eight measures submitted for consideration by the Steering Committee.

My name is Diedre Joseph, measure development staff for the AMA-PCPI. Also present to speak on behalf of the measures are Dr. Irene Katzan and Dr. David Seidenwurm, members of our Measure Development Work Group; Dr. Mark Antman, Director of Measure Development Operations at the AMA-PCPI; Kendra Hanley, member of the AMA-PCPI Specifications Team, and Laura Yodice, member of our Testing Team.

The American Academy of Neurology, American College of Radiology, and the American Medical Association convened PCPI, collaborated to form a Stroke and Stroke Rehabilitation Work Group in order to identify and define quality measures toward managing and improving outcomes for patients with stroke and for patients undergoing stroke rehabilitation.

The eight measures presented for your review were developed as a part of this measure development project. The measures were originally approved by the PCPI membership in 2006 and were recently reviewed and revised per the AMA-PCPI measure maintenance review and enhancement process. Six of the eight measures were originally endorsed by the NQF in 2007.

PCPI measures are developed through cross-specialty, multidisciplinary work groups. All medical specialties and other healthcare professional disciplines

participating in patient care for the clinical condition or topic under study must be equal contributors to the measure development process.

In addition, the PCPI strives to include on its Work Groups individuals representing the perspectives of patients, consumers, private health plans, and employers. This broad-based approach to measure development ensures buy-in on the measures from all stakeholders and minimizes bias toward any individual specialty or stakeholder group.

All Work Groups have at least two Co-Chairs who have relevant clinical and/or measure development expertise and who are responsible for ensuring that consensus is achieved and that all perspectives are voiced.

All eight of the measures presented today have been tested and shown to be reliable, feasible, and valid. The testing project for these measures included the assessment of face validity, inter-rater

reliability, feasibility, and for some a comparison across multiple data sources.

Face validity was systematically assessed by a panel of experts to establish the measure's ability to accurately reflect quality as specified. Additionally, all PCPI performance measures are assessed for content validity by a panel of expert Work Group members during the development process. Additional input on the content validity of draft measures is obtained through a 30-day public comment period and by also soliciting comments from a panel of consumer, purchaser, and patient representatives convened by the PCPI specifically for this purpose. All comments received are reviewed by the expert Work Group and the measures adjusted as needed.

Inter-rater reliability was assessed with a sample from several practice sites representing a range of settings, geographic locations, and technology infrastructure. The

percent agreement between trained extractors was recorded, as well as kappa statistics, to ensure that agreement rates are not a phenomenon of chance. All of the measures showed high agreement rates, and the kappas, when calculable, reflect substantial or almost perfect agreement beyond chance.

Finally, across the sites who are submitting data to PCQRS, four of the measures were tested for reliability across data sources. This allowed us to determine a percent agreement between PQRS claims submissions and a manual review of the claims information. The percent agreement for the four measures ranged from 91 percent to 100 percent agreement, reflecting high reliability in comparison of these data sources.

Now a brief review of the measures and rationale. There are eight of them, so please bear with me.

Measure 0242, Tissue Plasminogen Activator, t-PA, considered, is paired with

Measure 2022, t-PA-initiated, and both are specified at the facility level of measurement.

CO-CHAIR TIRSCHWELL: Sorry. I apologize for interrupting. But we are going to go through each measure in tremendous detail.

MS. JOSEPH: All right. So, we were told --

CO-CHAIR TIRSCHWELL: We only have about two minutes per organization.

MS. JOSEPH: We were told to give an overview of the testing and the methodology for measure development, as well as each measure.

CO-CHAIR KNOWLTON: Well, if the Committee has questions, we will direct them to you, but we have got to move on.

MS. JOSEPH: So, you don't want me to do each measure?

CO-CHAIR KNOWLTON: Right.

MS. JOSEPH: Okay.

CO-CHAIR TIRSCHWELL: We are going to do each measure.

MS. JOSEPH: Okay. Thanks.

CO-CHAIR TIRSCHWELL: Thank you.

MS. JOHNSON: Can we hear now from The Joint Commission developers?

MS. KOLBUSZ: Good morning, everyone.

My name is Karen Kolbusz. I am the Associate Project Director in the Division of Healthcare Quality Evaluation at The Joint Commission.

To my left I have Ann Watt, Associate Director, and also Dr. Lee Schwamm, who is a vascular neurologist and the Stroke Director and Director of Telemedicine at Massachusetts General Hospital, also Professor at Harvard Medical School, and Chief Consultant with Get With The Guidelines. Hopefully, I got all his titles correct.

(Laughter.)

To give you just a brief overview of the measure set, The Joint Commission stroke core

measure set consists of eight core measures.

The measure that deals with acute intervention is Stroke 04, which is thrombolytic therapy. This particular measure captures those patients who arrive in the hospital emergency department within two hours of last-known well that received IV t-PA within three hours of last-known well.

The next two measures are early management measures which involve VTE prophylaxis and also antithrombotic therapy by the end of day two, with day one being the hospital day of arrival, day two the day after arrival.

The remaining six measures deal really with discharge planning: patients discharged on antithrombotic therapy, patients with atrial fibrillation who are discharged on anticoagulation therapy, patients that are discharged on statin medication. These are all dealing with ischemic stroke patients. Also, patients that receive stroke education

or who are assessed for rehabilitation services anytime during the hospital stay.

The measures were originally pilot-tested in 2004. There was a 12-month pilot test. They have been in use since 2004. They were originally endorsed by NQF in 2008. Currently, they are collected by more than 900 certified Joint Commission primary stroke centers. Data collection for all stroke measures, all eight in the set, are required to maintain your certification status. They have also been harmonized with the Guidelines in the Paul Coverdell National Acute Stroke Registry.

They will be collected by CMS for the fiscal year 2015 payment determination. Therefore, data collection will begin in January of 2013.

And finally, they are included in the 15 clinical quality measures for Meaningful Use. That is a fiscal initiative also from the government.

Was there a question?

Okay, and that concludes. Thank you very much.

MS. JOHNSON: And now, if the folks from the American Heart Association/American Stroke Association would like to speak?

DR. SCHWAMM: Hi. I am Dr. Schwamm. You just heard my illustrious introduction from Karen. I am also here representing the American Heart Association.

I am here with my colleague, Mark Stewart, who is a staff member at AHA. Penelope Solis, also a staff member, is on the telephone. And Greg Fonarow, who is a cardiologist at UCLA and also a developer of measures within the American Heart Association in his capacity as a volunteer, may join by the telephone.

There are two measures before you from the American Heart Association. The first is timeliness of thrombolytic therapy. So, this is a measure that looks at the percentage of

patients among those receiving thrombolytic therapy who do so within 60 minutes, which is the recommendations from the NIH and the American Heart Association and all major guideline-issuing entities, and is supported by Level 1 evidence.

The measure has been in place since 2003 within Get With The Guidelines, which now has 1600 hospitals collecting data on this measure, and over 74,000 patients have been treated in the context of Get With The Guidelines and had this measure assessed.

And the measure also reports a median time to treatment. So, it supports the use of feedback to hospitals for improvement and helps identify hospitals that are near the boundary of that measure, but don't necessarily meet the 60-minute mark.

There is very strong data to support, both in meta-analyses and in our own work, that earlier treatment leads to better outcomes and reduced mortality. So, a strong

process/outcome link.

The second measure is the percentage of patients who receive an NIH stroke scale to assess their stroke severity on arrival. That measure has been assessed in over 800,000 patients, recorded in over 800,000, assessed in over 1.5 million patients.

It is a very important measure to identify the assessment of stroke severity at onset. It has strong linkages to both the use of thrombolytic therapy as well as access to comprehensive stroke center resources, and is strongly predictive of outcome, particularly in-hospital mortality.

I am happy to take any questions. Otherwise, thank you.

MS. JOHNSON: Okay. Thank you very much.

We will now begin our day with Measure 0437. This is the Stroke 04 measure from The Joint Commission, Thrombolytic Therapy.

We tried to set up this summary document

in the order in which we will go today. So, you should see that beginning on page 3.

And I am now turning it over to the Co-Chairs.

CO-CHAIR KNOWLTON: Just a reminder that we talked about earlier. As we call on you, just click on your microphone, please. Go through the Chair. Keeping this orderly, we will see that we get through it.

We are going to present each of the measures. We are going to ask the person who presented it in your Working Group to present it here.

But we are going to target the discussion, if you would, please, to the issues that we will be voting on. So, we will start with impact and we will work through the list. You will see it on the slides where it says "voting slides" in front of you. that will help us keep this organized.

So, for our first measure, which is Measure 0437, Thrombolytic Therapy, Greg, if

you would start off with that, please?

MEMBER KAPINOS: This measure was proposed by The Joint Commission for maintenance. We just heard the description, but, again, it is measuring, out of all the patients that present with an acute ischemic stroke within two hours from their last-known well to the ER, how many will, indeed, receive t-PA within the three hours. So, it gives us an hour in the ER to give t-PA.

So, the numerator, as we said, is the patients who receive t-PA within two hours, and the denominator is the ones that were eligible for this, which is within two hours of their stroke onset they reach the ER.

I think in terms of importance, the group did not have any more questions after the meeting over the phone. So, the impact was unanimously ranked as high for 1a.

1b, there were no questions --

CO-CHAIR KNOWLTON: Greg, stop at the impact. You discuss where you got an impact,

and then ask if there are any questions.

MEMBER KAPINOS: Okay.

CO-CHAIR KNOWLTON: And then, we will vote on it.

MEMBER KAPINOS: So, for 1a, impact, anybody?

CO-CHAIR TIRSCHWELL: You shared what the committee did, but this will be awkward until we get used to it. But this is the most efficient way; trust me.

So, does anybody have any questions on the impact that they would like to put to Greg or any other member of the project team that worked on this?

(No response.)

No questions?

Then, you are ready for a vote. I don't know how these clickers work, Karen.

MS. JOHNSON: We are going to have Suzanne or Jessica give us a little tutorial on how to use our clickers.

MS. THEBERGE: Okay. I am going to pull

up the vote and just click. For your rating, 1 is high, 2 is moderate, 3 is low, and 4 is insufficient evidence. So, just click and, as people vote, it starts to show up in that little box.

(Vote taken.)

We have got four responses, eight, et cetera. Seventeen. We are doing good.

MS. JOHNSON: And just a reminder, up on the screen are the two things that you would be thinking about under impact.

MS. THEBERGE: Twenty, and we need one -- oh, 22. Okay, and we are done.

MS. JOHNSON: Do the colors mean anything on the clicker?

MS. THEBERGE: No, the colors don't mean anything.

So, this has passed with 20 votes for high and 2 votes for moderate on importance.

CO-CHAIR TIRSCHWELL: In terms of impact.

MS. THEBERGE: Sorry. Impact.

CO-CHAIR TIRSCHWELL: Okay. Greg, go on to your next piece.

You see, the reason we do this way is, if we don't -- this is a must-pass -- if we did do the impact and it didn't pass, we would stop discussion and move on to the next measure. So, that is why we are doing it in order.

Go ahead, Greg.

MEMBER KAPINOS: So, for 1b, the performance gap, there was --

CO-CHAIR TIRSCHWELL: Actually, Greg, sorry to interrupt. We want you to do the evidence second and then we will go to 1b. So, 1c first.

CO-CHAIR KNOWLTON: Yes, they are in the wrong order on the slide.

MEMBER KAPINOS: Okay. For the evidence, five of us ranked it as high; two of us ranked it as medium. There were no major issues that were brought up during the meeting.

CO-CHAIR TIRSCHWELL: Any questions for Greg?

(No response.)

Okay. We are ready to vote.

MS. THEBERGE: And Jessica has just reminded me to tell you all that you can click more than once, and it helps to point your clickers towards me because the voting reception thing is over here. But it will only count your vote once.

(Vote taken.)

So, we are still waiting for several votes. All right, one more. All right. We are waiting for one more. It seems like somebody's click didn't come through.

CO-CHAIR KNOWLTON: Just vote again. It won't count you twice.

Is there anybody that didn't vote?

One more time.

MS. THEBERGE: Well, we had 22 votes last time.

CO-CHAIR TIRSCHWELL: There's 22 votes.

I counted them.

MS. THEBERGE: All right. So, the vote count is 18 yes and 3 no.

MEMBER KAPINOS: So, what happens when there are noes? It is just the majority wins? So, even if it is like borderline, like 16-15, you are going to make it pass?

CO-CHAIR TIRSCHWELL: I guess we had hoped there would be discussion if it was ambiguous like that.

I am curious why people voted no, but I don't think we have time to necessarily delve into it.

CO-CHAIR KNOWLTON: But I think what we want to do is just it is okay to say, "Look, I've read what is in the project sheet, and I understand people disagree with me," and move on. But it is also, if you feel strongly, I hope you will speak out.

Part of our job is to keep this moving, keep it going, but it is also to make sure everybody gets heard. So, it is very

important, if you think you have a persuasive argument, you should make that argument, and maybe somebody will agree with you.

Okay, Greg, let's move on to -- what are we on?

MEMBER KAPINOS: Performance gap, 1b? Right? Performance gap, 1b, we didn't do it.

So, there were a few questions that were raised there because of actually -- no, actually, it was something else. Yes, for this measure, the performance gap is there. Not a lot of hospitals are using t-PA, even for eligible patients. So, the majority of us voted that there is a performance gap. There are no specific questions that were raised, I think.

Ready to vote? Or do you have any questions?

CO-CHAIR KNOWLTON: Questions or comments?

(No response.)

Okay. Let's vote.

(Vote taken.)

MS. THEBERGE: All right. We are still at 21.

MS. JOHNSON: If everyone could go ahead and vote again?

MS. THEBERGE: Point at me, please.

All right. There we go.

CO-CHAIR KNOWLTON: Somebody has got a sticky clicker.

MS. THEBERGE: So, we have 21 votes for high and 1 for moderate.

CO-CHAIR KNOWLTON: Okay.

MEMBER KAPINOS: So, moving on to 2a?

CO-CHAIR TIRSCHWELL: Yes.

MEMBER KAPINOS: 2a, the reliability of this measure, I don't recall any specific issues there. Again, the majority of us voted for high quality for item 2a.

CO-CHAIR KNOWLTON: Any questions on this?

(No response.)

Okay. We can vote.

Hold on.

CO-CHAIR TIRSCHWELL: Can I just --

CO-CHAIR KNOWLTON: Sure.

CO-CHAIR TIRSCHWELL: There were a couple of things that were brought up during the Work Group discussion about the numerator and the denominator. The developer explained that the denominator -- and this is on the back of the page in the summary document. I guess there was a question about that two hours and the three hours, and it is just clarification that it gives the hospital at least 60 minutes to make the determination and begin treatment. I don't know that there are any questions about that.

And there was another question about why length of stay greater than 120 days was excluded. And that is just an artifact of billing cycles for CMS.

MEMBER J. BAUTISTA: I don't really understand that. What does that mean?

CO-CHAIR TIRSCHWELL: Which part?

MEMBER J. BAUTISTA: That 120-day --

CO-CHAIR TIRSCHWELL: The billing cycle?

MEMBER J. BAUTISTA: Yes.

CO-CHAIR TIRSCHWELL: I think it has to do with quarterly reports and billings. And so, if your stay extends over more than 120 days, and they do the previous 120 days, you can sort of be out of phase with the whole thing.

If I had a chalkboard, I could probably explain it better than that. But my guess is that it is just a really minor issue in general, and I wouldn't worry about it too much. I believe them, that it doesn't work very well with their computers.

CO-CHAIR KNOWLTON: That question has been asked a couple of times. Can I ask the developer to just clarify that?

MS. KOLBUSZ: The issue with regard to the 120?

CO-CHAIR KNOWLTON: Yes.

MS. KOLBUSZ: This is a set of measures,

one of many sets of measures that The Joint Commission has that we share with CMS. It is just a standard exclusion for all of those measures, that patients who are in the hospital for longer than 120 days are excluded because of the CMS billing cycle. Medicare requires patients to be billed every quarter. And the problem is, if they stay over 120 days and then are discharged in the next quarter, that history from the previous 120 days gets lost. So, anybody over 120 days is excluded from the measure. In real-life for this measure, well, as you saw in the submission materials, it is negligible.

CO-CHAIR KNOWLTON: Okay. Thank you.

Did that answer your question, Jocelyn?

Anything else?

(No response.)

Okay. Are we ready to vote on this?

We are voting first on reliability. There we go.

(Vote taken.)

MS. THEBERGE: We are at 19. So, we need three more votes. One more vote.

All right. There we go.

CO-CHAIR KNOWLTON: Vote early, vote often.

MS. THEBERGE: We have 15 at high, 6 at moderate, and 1 at low.

MS. JOHNSON: Can I just jump in here real quickly?

Can you put up slide 16? Just to remind you on how you rate reliability and validity, you should be thinking about levels of testing that were done. So, when you think about rating for reliability, for example, you are not looking only at the results of the reliability test, but you are looking at whether or not things were done at both the data element level and the measure score level or only one or the other.

So, for example, a high rating for reliability, you would expect to see precise specifications as well as testing and

demonstration of adequate results at both the data element level and the measure score level in order to give the measure a high rating. Moderate would be precise specifications, but only seeing testing at one or the other levels.

So, I just want to make sure everybody was clear about how the ratings for reliability and validity work. Validity is similar. So, again, the high and moderate ratings depend -- you have to be convinced that the results are adequate, but, also, it depends on how much testing has been done.

MEMBER KAPINOS: Maybe the people who voted moderate are thinking -- there was issue that was brought up during the call conference about comparing this measure to another one that had exactly the AHA guidelines inclusion and exclusion criteria for t-PA versus this measure, Stroke 04, by The Joint Commission is actually just including in the exclusion criteria a documented reason for not

initiating IV thrombolytic.

So, I don't know what is the rationale for the people who are voting moderate, but from that conference call I recall that people were comparing to the other measure, and maybe this one is actually simpler because we are just saying, as soon as you document why you are not giving t-PA, you are going to be excluded. Therefore, you are not going to be dinged by this measure.

So, just to open that to the discussion. If people who voted moderate are thinking about the fact that the exclusion criteria for t-PA are not thorough enough, there is actually another measure that does include all those exclusion criteria and to t-PA.

MS. JOHNSON: And just to clarify, and you may be right, but what you want to do is be voting on this measure, and this measure only. So, for now in this process, forget that there are other measures that are similar because we will be handling that a little bit

later in the day. That will be our related/competing.

So, yes, you would definitely only be thinking about this one, not how this one maybe stacks up to another one.

MEMBER COONEY: Since we are taking a little more care with this one, could we maybe look at this one specifically with that in mind, just to show us in the text of this where -- I mean, I see where the data elements are. I am not sure whether I see where the reliability of the measure score is in this one. That is just so that I can find it more quickly in others.

CO-CHAIR KNOWLTON: Yes, I think, Karen, can you walk them through that? Because I think you are right; I see --

MEMBER COONEY: I mean, you are making a point that it is important, and I am not sure I --

CO-CHAIR KNOWLTON: Yes, people aren't following it. I see Gwen nodding. So, let's

walk them through it.

MS. JOHNSON: And Suzanne is bringing up that measure right now.

I am sorry. I had to pull out my paper copy because I can't see that from here, even with glasses on.

Okay. So, for the reliability testing, if you have the measure submission in front of you, it should be somewhere around page 15, it looks like on that screen.

What they tell you, under 2a2.2, their analytic method, they tell you that they looked at 900 patient records from 30 different sites and, basically, did a re-abstraction. So, they compared data. So, they have the abstraction that was made for the measure, and then The Joint Commission abstractor also did it, and then they compared the results.

So, under Section 2a2.3, which is page -- is it 16? -- page 23, sorry, these are the results of their testing between the

abstractors. We do know that this isn't necessarily pretty, and this is not The Joint Commission's fault. This is because we can't accept pretty tables. So, they go in as just plain, unformatted text.

But what they have done here is they have given you several data elements, and what they are showing you is the numerator. So, how many records were looked at? And how many met the measure? So, if you go across there, clinical trial, it looks like they looked at the total "N", 712. So, what they are saying here is 712, of the 714 records that they looked at, 712 agreed on whether or not the patient was included in a clinical trial. And for this measure, clinical trial was one of the exclusions.

So, what they are showing you over to the right is the 99.7 percent, that is the percent agreement. So, 712 out of 714 is the 99 percent.

So, when you are doing this kind of

comparing between what one person says and what another person says, that is generally thought of as inter-rater reliability. And, of course, you want that to be high.

What they have shown you here is the statistic that is the percentage agreement. And that is a fine statistic. It is probably not the greatest statistic for showing you because, if you take two people off the street and ask them a question, often they will agree, even though they might not know anything of what they are talking about.

So, there is another statistic that they could have shown you called the kappa statistic, but they chose not to do that, and that is fine. But the kappa statistic often will be informative as well because that gets rid of that -- it takes into account chance agreement that could happen.

So, what they have done is given you many of the critical data elements. So, if you recall from our tutorials, if they are

going to give you data element reliability, they need to do it for the critical data elements. So, there are several elements that are critical for this measure. So, the things that go in the numerator, the things that are in the denominator, and the denominator exclusions.

So, as you work your way through those percentage agreements there, you would look at those and just ask yourself, you know, am I pretty convinced that this is reliable, based on this data? And you would just look at that and decide if you think those percentages there are good.

So, what you are doing is you are thinking about their methodology, which is the two abstractors. You are thinking about their method and their sample size. They told you it was roughly 900 records total over 30 sites. So, does that seem reasonable to you?

And then, you are looking at the results, the percentage agreements. Any

difficulties with that? Anything kind of raise a flag to you that you would think that there might not be some reliability there?

So, that is how you would check and think about reliability or at least the reliability testing. So, does that make sense?

MEMBER COONEY: Well, that, to me, looks like we are looking at data elements, but you said, also, that it had to look at measure score in order to be high?

MS. JOHNSON: To be high, yes.

MEMBER COONEY: So, does this look at measure score anywhere?

MS. JOHNSON: This does not look at reliability for the measure score. So, right here, what you are seeing is data element reliability.

What they did here as well with their validity, they can use these same results as their validity testing at the data element level as well. The reason they can do that is

because their second abstractor from The Joint Commission they consider the authoritative source.

So, if you agree that The Joint Commission abstractor could serve as an authoritative source, then we would also accept this result here as data element validity. So, you can check that box off, again, if you like those percentages.

And then, we would go further in the submission to see, would they also have done measure score validity or not?

So, is that starting to -- everybody is getting that?

MEMBER COONEY: So, sometime when they show measure score validity, will you point it out to us?

MS. JOHNSON: Yes.

CO-CHAIR KNOWLTON: Jocelyn?

MEMBER J. BAUTISTA: Would there be a reason why the sample size for the data element "time last-known well" would be so

much smaller than, for instance, clinical trial? Would there be some valid reason why that would be?

CO-CHAIR TIRSCHWELL: Well, yes. So, I asked this, I think, during our Work Group call. I think it has to do with the number of people who would qualify. So, these aren't patients that just got IV t-PA. It is a whole bunch of people that didn't, for reasons. And so, who gets the different elements is dependent on that, I think.

You know, and I would just add that -- and I think this is true to clinical practice -- the time last-known well, there was only 80 percent agreement. And I think we, all of clinicians, can quote you cases where it gets pretty murky exactly when the last time known well was. And a couple of the lower agreement rates were why I rated it moderate, as a comment.

CO-CHAIR KNOWLTON: So, on that same point, what I am not understanding about this

-- and this is true in some other things -- the different numbers that are being given for the different data elements, I made the assumption that the smaller numbers were subsets of the larger number. Because there is a large group of patients, and then, within that, there is a smaller subset that meets whatever that criteria is.

What I don't understand is, why, then, is the larger number relevant? Let's say here you have got, whatever it is, 198 patients or something that meet the criteria of time from last-known well, when that is essential to this particular measure, right, because there is, clearly, a lot of discussion about that point in your group, or whatever, if that is essential to it, then why is the broader number that doesn't include that relevant? That is what I don't understand.

Like, why are we seeing numbers that say the overall agreement rate is 98 percent, and we are seeing 99.7 percent in sort of clinical

trials of like whatever the overall number is. But the most relevant numbers, the ones that relate specifically to this measure, are lower numbers. Why are we even being given those broader numbers?

CO-CHAIR TIRSCHWELL: Let me see if I am getting this right. On the top line there, they say 98.1 percent. I am guessing that maybe is just an average of this entire set of columns.

CO-CHAIR KNOWLTON: Right, but I am trying to understand what the relevance is of all the different parts of this column.

CO-CHAIR TIRSCHWELL: Right.

CO-CHAIR KNOWLTON: Because it seems to me -- and maybe I am misunderstanding this, and maybe they can clarify it for us -- but it seems to me that, if there are these subsets -- because this relates to how we understand all of the other measures going forward.

CO-CHAIR TIRSCHWELL: Sure.

CO-CHAIR KNOWLTON: That is why I think

it is worth understanding now.

CO-CHAIR TIRSCHWELL: Well, yes, and specifically, The Joint Commission measures because they all present the data like this.

CO-CHAIR KNOWLTON: They use a lot of this type of thing.

CO-CHAIR TIRSCHWELL: So, could we ask the developers to comment, please?

DR. SCHWAMM: So, I think maybe this will help. All of the elements that are listed there are part of the measure construct. Because, for example, if you are in a clinical trial, that is an exclusion to the measure.

So, I think the appropriate way to report the reliability is to review every element that must be abstracted responsibly and reliably in order for the construct to be valid.

So, you are correct, I think, in identifying that there is better agreement around whether someone was in a clinical trial

than the precise documentation of the time last seen well. But if you didn't present that number at all, and people were not reliably abstracting the presence of a clinical trial, the measure construct could be invalid and you would never know it.

So, you are focused on the numerator/denominator statement around the numerator. But if you think about the denominator, you have to also accurately abstract the exclusions from the denominator. So, every element listed there is part of the measure. And if you reviewed the measure description in detail, you would see that each one of those has a specified variable, and that there are guidelines to the abstractors for whether or not it counts as present or absent.

Is that helpful?

MEMBER KAPLITT: I think that is very helpful. I do think that it is a little bit -- "misleading" is probably too strong a term

because I am not suggesting any overt intent, but I think that the problem is the overall statement at the top, which is a recurrent theme in a lot of The Joint Commission measures, about the overall reliability rating, I think gives a false impression, whether it is intended or not. And I assume it is not.

But I understand your point. I mean, I think maybe going forward, not that we are going to change everything now, but it might be more helpful in the future to sort of categorize the different submeasures. Because I understand your point, but they are not all necessarily equal or equally-relevant. And they are presented, when they are given as the overall measure, as if they were all kind of an average of one thing.

DR. SCHWAMM: Yes. And actually, as a clinician, I totally agree with you, but wearing kind of more of a measurement hat, I think we place a lot of emphasis on what we

consider to be the critical data element. But abstractors have to actually figure these things out.

And let's say, for example, being on a statin was an exclusion to the measure. Well, if there is data that shows that the abstractors don't accurately identify drugs that are statins and fail to do that accurately, then you need to know that.

So, I think it is reassuring --

MEMBER KAPLITT: No, I take your point, clearly.

DR. SCHWAMM: Yes.

MEMBER KAPLITT: And I think most of us get that at this point.

DR. SCHWAMM: Yes.

MEMBER KAPLITT: All I am saying is that the flip side is also an issue, right?

DR. SCHWAMM: Yes.

MEMBER KAPLITT: So, yes, you are right. It is like the way we are going through all these different criteria, one, two, three,

right?

DR. SCHWAMM: Exactly.

MEMBER KAPLITT: You have got to pass one, but --

DR. SCHWAMM: Exactly.

MEMBER KAPLITT: -- that doesn't necessarily mean that that is sufficient. So, being necessary is one thing, but being sufficient is another thing. And I think that that is where the confusion --

DR. SCHWAMM: Completely agree, and I think that is why you get all the individual variables listed there.

CO-CHAIR KNOWLTON: Let me ask the question, since we have had deeper discussion of this, do you want to re-vote this issue?

Okay. Can we do that, Suzanne?

This will be on 1b?

CO-CHAIR TIRSCHWELL: No.

CO-CHAIR KNOWLTON: What are we on?

CO-CHAIR TIRSCHWELL: 2a.

CO-CHAIR KNOWLTON: 2a.

She has got to open it. There you go.

MS. THEBERGE: Okay, it is open.

(Vote retaken.)

Nineteen. Twenty-one. We are still waiting for one vote. Oh, there we go.

So, we have 4 high --

CO-CHAIR KNOWLTON: Teachable, or we are all cowards.

(Laughter.)

MS. THEBERGE: Four high, 17 moderate, 1 low.

CO-CHAIR KNOWLTON: Okay, Greg, we go on to validity.

MEMBER KAPINOS: Yes. For criteria of validity, again, there were no major issues and unanimously we voted that it was a valid measure.

CO-CHAIR KNOWLTON: Any questions on validity for Greg?

(No response.)

Okay. Let's vote on validity, 2b.

MS. THEBERGE: Sorry. There we go.

Oops. Sorry, we are having a little technical glitch here.

CO-CHAIR KNOWLTON: There you go.

Yes, she has got to open it. Wait.

No, not yet.

MS. THEBERGE: There seems to be a technical --

CO-CHAIR KNOWLTON: There you go.

MS. THEBERGE: All right. There we go. Sorry about that.

(Vote taken.)

Eighteen. Nineteen. Twenty-one. All right, we need one more. All right.

Thirteen high, 8 moderate, and 1 low.

CO-CHAIR KNOWLTON: Okay. Okay. We are moving on to --

MEMBER KAPINOS: Usability.

CO-CHAIR KNOWLTON: I'm sorry?

MEMBER KAPINOS: Usability, item 3.

CO-CHAIR KNOWLTON: Okay.

MEMBER KAPINOS: The Work Group thought that there was high evidence of the usability

of this measure. No specific question.

CO-CHAIR KNOWLTON: No specific discussion?

Any specific discussion within the group? Dave?

CO-CHAIR TIRSCHWELL: You know, it says that it is used for public reporting. I guess I don't know where this is publicly-reported. Or is it just publicly-reported, not at the hospital level? Can the developer comment on that?

MS. KOLBUSZ: The public reporting refers to CMS's use in the fiscal year 2015 payment determination. It was determined in the IPPS final rule last year, in Quality Check as well, and also in Quality Check. But the data -- let me finish.

The data will collected by CMS in 2013, January 1st. Also, it is publicly-reported by The Joint Commission and Quality Check. That has been ongoing since it has been made a core measure in 2009. So, it is publicly-reported

on The Joint Commission website and Quality Check.

CO-CHAIR KNOWLTON: Go ahead. Risha?

MEMBER GIDWANI: Thank you.

Can the developers confirm that the CMS Meaningful Use specifications are 100 percent the same as their specifications?

MS. KOLBUSZ: Well, that is a very interesting question. I can confirm that the measures, the paper-based measures, are exactly the same.

If you are asking me, is there an exact one-to-one correspondence between the e-measures derived from the paper-based measures and the paper-based measures, the answer to that is unknown. Those measures, the Meaningful Use e-measures have not been tested. But that is certainly the intent.

CO-CHAIR KNOWLTON: Risha, would you comment on what is behind your question? Where are you going?

MEMBER GIDWANI: Well, I just want to

make sure that, if organizations are collecting these data for CMS Meaningful Use, they can submit the exact same data to The Joint Commission, and that it is not a duplication of effort for the institution.

MS. KOLBUSZ: Could I comment on that, please?

We, The Joint Commission, have developed a policy whereby hospitals who wish to report these measures for Meaningful Use can report either the e-measures or the paper-based measures. There is still a wide variation across the country in terms of the adoption of EHRs. Not all hospitals, in fact, the majority of hospitals don't have the capability yet of reporting on the e-measures. So, we are accepting both. We are not requiring both. I guess we are accepting either and not requiring that they do both.

CO-CHAIR KNOWLTON: Any other questions or comments on usability?

This is an important criteria,

especially as purchasers and consumers start to look at this data. So, I want to be sure everybody understands the criteria that is up there, as you look at what is usability.

Do we vote on these separately, Suzanne?

MS. THEBERGE: No, usability by itself.

CO-CHAIR KNOWLTON: Usability by itself. So, the two together?

MS. THEBERGE: Uh-hum.

CO-CHAIR TIRSCHWELL: A and B.

CO-CHAIR KNOWLTON: A and B?

MS. THEBERGE: Yes.

CO-CHAIR KNOWLTON: Any questions?

(No response.)

Okay. Yes, go ahead, Michael. I didn't see you.

MEMBER KAPLITT: No, because I didn't do it.

(Laughter.)

The one statement here, it says, since 2009 -- maybe the developer can answer this -- since 2009, it says, "Aggregate performance

has increased from 48 percent to nearly 72 percent."

Has the number of patients in the whatever, the denominator, or the number of patients being reported on been relatively-flat or increased during that time? Or does that percent reflect some sort of a shift in who is being reported on? Do we know?

DR. SCHWAMM: Let me answer that wearing both hats because almost all Joint Commission hospitals use Get With The Guidelines to report their data. I can tell you we looked at this, and this is published in a paper with Matt Reeves as the first author, looking at whether increased improvement was due to increased denominator exclusions or an actual increase in the rate of treated patients. And the denominator exclusions appear flat, and the rate of actual treatment appears to be increasing in a statistically-significant manner.

So, I believe that in the very beginning

of the program, in the sort of early 2000s, there was a shift with increasing denominator exclusions, but that has actually been very flat.

CO-CHAIR KNOWLTON: Any other questions?

(No response.)

So, we are looking at whether the usability criteria is high, moderate, low, or insufficient information. And we can vote.

Let's try it again.

MS. THEBERGE: Sorry about this.

(Vote taken.)

Nineteen. Twenty-one again. We are still short one vote. Can everyone vote again? All right. There we go.

Sixteen high, 4 moderate, 2 low.

CO-CHAIR KNOWLTON: Okay. Feasibility, Greg.

MEMBER KAPINOS: Looking at feasibility, we just addressed the concern about electronic medical record versus paper record, but, otherwise, there was an unanimous decision to

qualify the evidence as high for feasibility for this measure.

CO-CHAIR KNOWLTON: Any questions?

(No response.)

I do think we discussed this. Go ahead. We can vote on this.

CO-CHAIR TIRSCHWELL: So, it is not open until you see that timer go on, I think, on the bottom right there.

(Vote taken.)

MS. THEBERGE: All right. We are at 21. All right.

It is 18 high, 3 moderate, and 1 low.

CO-CHAIR KNOWLTON: And then, finally, the overall suitability for endorsement.

MS. THEBERGE: Do you want to vote now?

CO-CHAIR KNOWLTON: Anybody have any comment on this before we vote on it, the final criteria?

(No response.)

Okay. We can vote on it.

Wait for her to open it. That is what

is bouncing us out. There you go.

(Vote taken.)

MS. THEBERGE: All right. Twenty-two yes, zero no.

CO-CHAIR KNOWLTON: Okay.

MEMBER KAPLITT: Sorry, I don't want to be a stickler. Can I just ask one quick process question? I am not trying to call anybody out. But somebody repeatedly voted low and then everybody said yes. I am just wondering, is that like an error? Because, by the criteria I heard, if there is a person who felt each time that it was low -- well, I guess that is possible, yes.

CO-CHAIR KNOWLTON: And medium is good enough.

MEMBER KAPLITT: No, I know. I am not questioning that. It is just that somebody was repeatedly voting low. I assume it was one person, but maybe it wasn't. I don't know.

DR. BURSTIN: And just so you know, we

have actually attached your name to your clicker. So, those are the kinds of things we can do post-talk and see, in fact, there are any strange lack of face validity there.

CO-CHAIR KNOWLTON: We will have a man visit your home after you leave.

(Laughter.)

DR. BURSTIN: It is actually very helpful.

CO-CHAIR KNOWLTON: We will just be giving the names of the low voters to the developers.

(Laughter.)

I am from New Jersey. We handle these things very directly.

(Laughter.)

CO-CHAIR TIRSCHWELL: Okay. So --

CO-CHAIR KNOWLTON: Just one more second, David.

Before we go on to the next one, we will go very quickly, but like Michael's issue, does anybody have any process issues as we

went through this?

(No response.)

This will move along. Debate will be more intense as the day goes on, I promise you. If you have been through this a few times, you will see how this grows. But it is a little bit mind-boggling as you start out. So, you should expect a little confusion as you start out.

Okay, David.

CO-CHAIR TIRSCHWELL: So, the next measure is No. 0242, Tissue Plasminogen Activator Considered, from the AMA-PCPI.

Bill, do you want to get us started?

MEMBER BARSAN: Great. Yes. Yes, just briefly, there was a lot of discussion about this at the Work Group.

So, the description is all patients with ischemic stroke who arrive at a hospital within 4.5 hours of last-known well who were considered for t-PA administration. So, consideration for t-PA administration is the

numerator. The denominator is all those patients who arrive at the hospital within 4.5 hours or less known well.

Do you want me to summarize some of the issues first or no? I'm sorry.

We will start with impact, fine. Okay. Good.

So, impact, you know, it was pretty unanimous that we felt that -- seven people voted yes as far as the impact, and most of those were high.

CO-CHAIR TIRSCHWELL: So, that is 1a.

Anybody, anything to discuss on 1a before we go to the vote?

(No response.)

Okay.

MEMBER KAPINOS: The discussion that you were about to go over was probably the reason for maybe 1 percent voting a little bit lower quality. So, we need to talk about the fact that people were disturbed maybe by the fact that measuring "considered" as opposed to

"given" is something that we need to measure.

MEMBER BARSAN: Well, yes, I think there were a couple of issues. One was that whole issue of what is "considering". Is consideration, I mean is that a documentation measure or is that really a process measure? I mean, what is that?

MEMBER J. BAUTISTA: Well, it should be specified. So, what does the measure specify? They define "considered," don't they? So, could you summarize for us how they define it?

MEMBER BARSAN: "Considered" means if you didn't list a reason for excluding, essentially.

CO-CHAIR TIRSCHWELL: You know, I think it means you had a comment on it, basically. If they came in, you had to say either why they didn't get it or, if they got it, obviously, you considered it.

MEMBER WADDY: So, can we find out why the measure was developed that way instead of actual?

CO-CHAIR TIRSCHWELL: So, the question going to the developer, why consider it as opposed to used?

DR. KATZAN: Hi. So, this was because it goes out to the 4.5-hour window. So, IV t-PA within three hours has been around for a while. The 3-to-4.5-hour period is something that is approved by the guidelines, is recommended by the guidelines, but it is not FDA-approved. And there are more exclusions for this measure.

So, the thought was that we wanted to make sure that people consider the use of IV t-PA out to the 4.5-hour window, this 3-to-4.5-hour window period. But we realize that there might be some additional factors to consider when making this decision out to that 4.5-hour period.

So that, this is a combined measure. This should be paired. The intent was to pair it with the measure that actually talks about actual administration of IV t-PA out to three

hours, but that we want there to be some quality measurements that go out to the 4.5-hour period.

MEMBER BARSAN: David, there is a little confusion on my part because I am not sure where some of these concerns that I have fit in terms of specifically saying this is impact, this is evidence, this is whatever.

So, a couple of things. One is IV t-PA is not FDA-approved for administration out to 4.5 hours. I think there's a lot of different ways this has been picked up at different places. Some places do it routinely; some places don't. Certainly, the criteria are different. That is one.

Two, I think the whole issue of whether or not this is just -- you know, does this really reflect a process or does this just reflect documentation? In other words, is there really any evidence that, if somebody checks off a list and they are saying, "Oh, yeah, I considered it and it is not good,"

does that really improve the process? And I don't know that there is any data for that.

And I guess the third thing is that this takes patients who arrive at the hospital within 4.5 hours of last-known well. They are never going to be a candidate at 4.5 hours or 4 hours or probably anywhere beyond 3.5 hours. So, it seems kind of silly to throw in however many patients may be in that indicator that will never fit the criteria and just automatically X'ed off because they get there too late.

I don't know where to put those, but those are, I think, the three concerns at the Work Group that were brought up most prominently.

CO-CHAIR TIRSCHWELL: Yes. Thank you.

And I agree, it is sometimes a little ambiguous, and maybe they overlap the categories.

I am being informed that on the back side where it says, "Work Group Call Summary,"

the NQF staff has kind of put the things that were brought up on the calls into their appropriate categories. This one, the 3-to-4.5-hour time window has been into the scientific acceptability, which, Karen, you are saying is Category 2, validity or reliability?

MS. JOHNSON: It actually can come up on both.

CO-CHAIR TIRSCHWELL: Okay.

MS. JOHNSON: It could be under reliability because it would reflect the specifications themselves. But it could also come up under validity because it is the marriage, if you will, between specifications and the evidence. So, it kind of hits both of those criteria.

CO-CHAIR TIRSCHWELL: Well, you know, impact probably is simpler than all of that. Is this addressing a health concern, the use of t-PA, that potentially has impact? I guess you could, again, undermine it a little bit by

saying, "Oh, what about that 3-to-4.5-hour time window?"

MEMBER BARSAN: Well, that was my confusion because impact, you know, anything that could lead to an increased use of t-PA is a huge impact. I mean, there is great evidence that that is a good thing. It is just I am not sure whether this one does it, and I am not sure where the issues fit in.

CO-CHAIR TIRSCHWELL: And if I could just ask for one more clarification, Dr. Katzan, I guess you suggested by your comments that it was specifically designed this way to focus on the 3-to-4.5-hour time window.

And then, before you answer, there was a question that went around that is a little more general about FDA approval and the ability to endorse measures that include aspects that are not FDA-approved. And we got a little bit of information back. Can somebody clarify that?

MS. JOHNSON: Helen, I think that would

be yours.

DR. BURSTIN: So, in general, we obviously have tried to have committees stay within whatever is the evidence and whatever the guidance is. And certainly, FDA offers guidance in this respect.

The way we drafted it in our Evidence Task Force Report was really more about the indications. I don't think we got to this level of specificity about time window, for example. So, I think this is really where the expert opinion of the Committee and the guidelines really come into play.

And I do think that the issue of whether "considered" is important can fit many places, but I actually do see it somewhat under impact as well, because you are talking about the impact of this particular measure, not just t-PA and stroke care. This is specific to this measure. So, I think the "considered" issue is fair game here. I think the time window is also fair game under evidence.

CO-CHAIR TIRSCHWELL: Yes.

Dr. Katzan, do you have anything to add?

MS. JOSEPH: So, before Dr. Katzan responds, I just wanted to clarify a couple of things. Really quickly, the FDA time window is approved for three hours for t-PA administration. And the measure that you are reviewing now, Measure 0242, is the consideration. It is not the administration of t-PA. The administration of t-PA is actually covered in the next measure, which is Measure 2022. That is within the time window of three hours for administering the t-PA.

So, the window for 3 to 4.5 hours is included in the t-PA considered measure because we are aware that there are existing hospitals that do administer t-PA up to 4.5 hours, and our measures do not encourage or capture anyone that does administer t-PA up to the 4.5-hour time window. The measures only capture the consideration of t-PA up to the 4.5-hour time window. And the measure pair,

the Measure 2022, actually covers patients that receive t-PA up to three hours. So, I just wanted to make that clarification.

DR. KATZAN: Yes, so we wanted to measure -- that actually goes out to the 4.5-hour window. So, that we didn't want to stop at three hours. I don't know whether focusing within that time window was the purpose of this measure, but we wanted to make sure to include the out-to-4.5-hour time window.

MEMBER R. BAUTISTA: Hello.

Yes, the AMA actually is a quality measure. Just to be considered for t-PA by itself really means not very much unless you have a reason for that.

So, in many ways, you can think that it might be a copout for some centers who don't give t-PA and say, "Because you were four hours and 15 minutes late," you know, something like that. So, in other words, consideration without reason I think is a less-than-perfect measure.

CO-CHAIR TIRSCHWELL: Right. So, I think that is the point. Again, this is just a documentation issue. And is this really going to drive care? If a hospital has made a decision not to give t-PA in that timeframe, do they really need to write it on every chart to fulfill this performance measure? So, I think that is where some of the doubts are creeping into this assessment.

Does anybody else have any other comments on impact? I think we should go ahead and vote, otherwise.

Bill, do you have one last one?

MEMBER BARSAN: The only last comment on impact is that, if you look at the impact of treating a patient within an hour and a half compared to three hours compared to four and a half hours, a huge impact in an hour and a half; a big impact, three hours. If you get out between 3 and 4.5 hours, statistically-significant, but much, much less. So, I mean, just less of an impact than just in terms of

the number of patients effective.

CO-CHAIR TIRSCHWELL: Jolynn?

MEMBER SUKO: This gets to the documentation issue. I guess, is there any evidence that creating checklists enables t-PA to be -- having those forcing functions in your systems or those kinds of things, does that actually drive care? And do we have any evidence anywhere? And has the measure developer considered do we have any data around that?

MS. JOSEPH: So, actually, the literature that we have included focuses on the lack of patients receiving t-PA. The performance gap based on the inclusion of the measure in the Physician Quality Reporting System, PQRS, shows that there is a gap in care of 78.8 percent, and the medical literature also indicates that, while t-PA use increases stroke recovery by up to 50 percent with a low serious complication rate, only 3 to 8.5 percent of potentially-eligible

patients receive the t-PA. So, that is the purpose of encouraging the actual consideration.

CO-CHAIR TIRSCHWELL: So, I think the answer to your question is, no, there is no literature supporting it.

Okay. I suggest that we go ahead and vote on impact. It has to be high or moderate to move forward.

(Vote taken.)

MS. THEBERGE: We are at 21. All right.

Two high, 8 moderate, 10 low, 2 insufficient evidence.

CO-CHAIR TIRSCHWELL: Hung jury.

So, 12 did not reach a moderate or high level; 10 did. So, at this point, further consideration of this measure will not be performed.

Okay. So, I think at this point we are thinking maybe we would take our break. Or do you want to push on to do one more? What do you think?

All right. That one went pretty fast, shockingly.

(Laughter.)

So, let's make Dr. Barsan do another one, 2022, t-PA initiated. This was the paired measure, again from the AMA-PCPI.

MEMBER BARSAN: Okay. Great.

So, this is 2022. And the numerator for this is patients for whom t-PA was initiated within three hours of last-known well, and the denominator, diagnosis of ischemic stroke who present within two hours of time last-known well who are eligible for t-PA.

It is a little bit -- I don't know if you want to compare it with the other one -- but that little phrase about who are eligible for t-PA makes it a little bit different than the previous one we discussed because there were exclusions for that in there.

Let me just go on down.

So, in terms of the impact, it was not unanimous in the group as far as the impact,

but, mostly, six voted high, one medium.

CO-CHAIR TIRSCHWELL: Any further discussion about impact for this one? And note that this one is confined to the 3-hour time window.

CO-CHAIR KNOWLTON: I have a question. Are these the only contraindications that are listed that would be considered? The exclusions, it says that "contraindication conditions that might lead to," et cetera, et cetera. They are spelled out in this measure. Is that because those are the only ones that they consider for this measure?

MEMBER BARSAN: So, we had a lot of discussion about that, because some of these are actual exclusions and some of those, if you look at the FDA labeling, it is taken right from the FDA labeling. Some of those things like advanced age, severity of stroke, and whatever, are things where one should give more consideration. In other words, there may be reasons not to treat, although there is

really, frankly, nothing in the NINDS t-PA which suggests you shouldn't treat somebody who is at an elderly age with a bad stroke. But it was in the FDA warning.

And so, you know, supposedly, the other category, or the other measure, although these things weren't explicitly listed, supposedly they actually are part of that measure. And documentation of those items could be listed as an exclusion, is what I was told, what we were told.

But these are just listed more explicitly. I guess I have an issue with that because I am not sure there is evidence that one shouldn't be treating patients that fit some of these criteria.

CO-CHAIR TIRSCHWELL: Right. I think these are suggested as valid reasons why you might not give it. I don't think that, if you gave it, it would necessarily be counted against you if you don't believe in them specifically.

They do comment that it was harmonized with The Joint Commission measure specifically, although if you look at The Joint Commission measure that we just looked at, it doesn't look exactly the same. They just say that you have listed a contraindication, which is simpler to think about for sure. So, there are definitely some differences there.

Any other comments about --

CO-CHAIR KNOWLTON: Just a further comment on that. Some of the things just struck me in the exclusions as, if this became a checklist for providers, some of these are troubling. I don't know in terms of a quality measure. They may be a rational measure from a payer perspective, but from a quality perspective, advanced stage and some of the other issues I thought were a bit much.

But you are the doctor; I am not. So, tell me what your clinical judgment is on some of these.

MEMBER BARSAN: Right. Well, I am a doctor, and I have a problem with it.

(Laughter.)

No, my big problem is that explicitly listing them I think almost gives people a rationale for not using it. "Oh, yeah, somebody is old. I can't use it." Or "Gee, they have got an NIHSS greater than 22. So, I had better not use t-PA."

And that really bothers me because those aren't legitimate reasons for not using it.

CO-CHAIR TIRSCHWELL: You know, I would only add that, at least in my center and almost every other center I have come across, checklists do exist for inclusion and exclusion criteria. And even though you might not agree with any one particular one, and there is always room for local clinical judgment, it looks a lot like this.

I guess my assessment of whether this particular list is going to have a huge, specific impact at the patient level, I am

pretty dubious that it would.

And you know, these exact lists are in the AHA/ASA guidelines. There is nothing, I think, that is too terribly surprising or new.

MEMBER WADDY: The one that I found really surprising was just the diabetic retinopathy. I mean, you know, if a patient comes in aphasic and they don't tell you they have vision problems, I mean, in the throes of a t-PA call, I don't think there are many people that are doing detailed fundoscopic exams for that.

MS. JOSEPH: Excuse me. Are we able to respond regarding the exceptions?

CO-CHAIR TIRSCHWELL: Sure. The developers want to respond briefly.

MS. JOSEPH: Thank you.

So, just for clarification, the list that you are looking at in the denominator exclusions box are actually denominator exceptions. We distinguish between the two. Denominator exclusions are actually absolutely

never to be included in the measure. Denominator exceptions allow for clinical judgment by the physician.

And these are actually listed only as examples. We always include examples in the parentheses there where it says, "e.g., contraindications and conditions that might lead to increased risk of bleeding or unfavorable outcomes". It is not meant to be an exhaustive list by any means, but we did include this list of contraindications and warnings because during the Work Group discussion they thought that this list was really important, especially since The Joint Commission measure is the measure that we have harmonized with, and they have also included this list which comes from the FDA label.

So, again, these are only meant to be a list of examples for medical exceptions. They are not exclusions from the measure.

CO-CHAIR TIRSCHWELL: Ramon?

MEMBER R. BAUTISTA: Yes. Would it make

more sense, then, to risk-stratify the measure rather than actually having a long exclusion criteria list? In other words, if you are comparing quality across different healthcare organizations, it would seem to make more sense to risk-stratify the measure rather than to say, you know -- because someplace, for example, if you take innercity folks with a higher rate of stroke-related drug problems, where you might not give t-PA, for example, then you are actually dinged for not doing that, given a measure like this.

But if you actually risk-stratify and say, okay, we are going to compare organizations across the country, basically, and the individual differences that they might have.

CO-CHAIR TIRSCHWELL: Yes, go ahead.

DR. BURSTIN: Just a question for the developer. This is actually specified as a facility-level measure, which is a bit unusual for the PCPI measures. And the exception

reporting, my understanding, is to allow the clinician to indicate that they have a particular exception that may not be captured otherwise in the measure, that they should be excepted from the measure.

That doesn't sort of work. I can't understand how that works for a facility to have that clinical judgment. Can you explain that? It is a little confusing.

MS. JOSEPH: So, yes, I see how that could be confusing. That is a good question.

I think the intent here was to include the exceptions, also, because the physician will still be evaluating the patient. I am not sure who I am speaking to. The physician will still be evaluating the patient. Any reason that the physician chose not to give the t-PA should still be documented. And so, even though it is a facility-level measure, we did opt to leave the medical exceptions in the measure because it was important to capture the reason for not administering the t-PA.

CO-CHAIR TIRSCHWELL: I think they just end up averaging it over all the cases at the facility and allow them to have a medical reason why they don't give it. I mean, theoretically, this should be identical to The Joint Commission measure. I guess The Joint Commission measure had the good fortune of not listing their specific exclusions, although they are theoretically the exact same thing.

So, I guess, again, I am not sure how much we need to get hung up on the individual details when physicians still have exception and discretion to give this to who they think are the appropriate patients.

And I don't think you would be dinged, by the way, if you wrote "This patient has drug abuse and I think for X, Y, or Z reasons there is an increased risk of hemorrhage. So, I am not giving it." That won't count against you.

Go ahead.

MEMBER R. BAUTISTA: All I am saying is

being a quality measure where you actually compare organizations across the board, people aren't going to know that. They are going to say that in Hospital X there is less t-PA being given compared to Hospital Y. And that is all they know. That is all the public knows without the reasons, but they won't understand anyway. That becomes a problem.

What I am saying is, if you have a measure that actually risk-stratifies the risk across different centers, you are able to at least ameliorate that problem.

CO-CHAIR TIRSCHWELL: You know, I think the number that would be presented publicly is the percentage of eligible patients that get it. And I don't know whether the numerator and the denominator presented so that they could figure out what percentage of their total ischemic population is getting it. But that is not the relevant number.

MEMBER BARSAN: That is part of my problem, is the way it is defined. So, it is

the number of potentially-eligible patients. That is the denominator, right?

I would be okay -- I mean, so how do you determine the number of potentially-eligible patients? That is the confusion. I mean, it seems to me like one way to do it is if you just said anybody who meets those exclusions from the FDA warning label, the exclusions only, that they are out. So, they are ineligible. Everybody else is eligible. And then, the only other ones, if for some reason you decided not to treat and you documented why, or whatever, that was okay.

But I don't know. It is just unclear to me.

Irene, can you help me with that? Do you see where my issue is? I am just not quite sure I understand how do you determine the patients for the denominator. I mean, how is this different than just saying everybody who presents within less than two hours?

DR. KATZAN: And I actually struggle

with the differences between The Joint Commission measure and this measure, because I think, operationally, I am in stroke neurology. I see patients and I treat them with IV t-PA. If they have an exclusion according to the guidelines or the FDA, we put them as an exclusion. They are excluded from the denominator. We don't actually include them in the measure.

And so, I think this is a lot of very detailed specifics, but I think the purpose and the rationale for including these measures was to help the people who are actually administering the t-PA, the people on the ground floor, to understand that these are the generally-accepted reasons for not administering IV t-PA in patients who come within this time window.

I think that, operationally, what we do, as stroke neurologists, they come in; we evaluate them. We actually have checklists; most of us do. If they fall and they don't

actually have an exclusion, then we administer t-PA, and if they do have an exclusion, we make sure that we document why or what the exclusion is.

And we use generally a list of things, and this is actually just a list that has been generally-accepted. It is on the FDA label. It is in the guidelines.

So, I actually don't see some of the issues because I think, operationally, this is how we practice.

MS. JOSEPH: I would just add briefly, to answer your question with regard to how to identify the patients in the denominator, the denominator statement is all patients age 18 years and older with a diagnosis of ischemic stroke who present within two hours of time last-known well and who are eligible for t-PA.

However, we also include in the denominator details, a definition for eligible. Patients are eligible for t-PA if they have an acute neurologic deficit, a

clearly-defined time of onset of less than 180 minutes before treatment, and a baseline CT showing no evidence of intracranial hemorrhage.

CO-CHAIR TIRSCHWELL: And also, implicitly, don't have an exclusion criteria --

MS. JOSEPH: Exactly.

CO-CHAIR TIRSCHWELL: -- which is, of course, coming back to what we are all struggling with.

Salina?

MEMBER WADDY: So, I understand how you would handle the contraindications. But if you have this measure as-is, how would you handle the warnings and conditions? Because that could be as many as 20 percent, 30 percent of the patients.

CO-CHAIR TIRSCHWELL: I think I can answer that. Basically, if the physician documents almost any reason why they feel it is not appropriate to give it, the patient is

put aside and not included in the measure.

This and I think The Joint Commission measure are both aimed at identifying patients. For example, at my center we goofed, and we had a good reason why we didn't give it, but we didn't document it. So, it reminds us that we have to document it.

And more importantly, at medical centers which don't have as much a commitment to stroke care, where somebody comes in and could have been a candidate, not only do they not even think about giving it and not document it, but, you know, those are the types of patients that we want to be called attention to, so that the appropriate patients are being considered.

So, I think if you are documenting your reasons, as I think any of the vascular neurologists at this table involved in National Quality Measures are probably doing pretty reliably, it won't be a problem for you. It is this other lower-quality care

scenario that is relevant.

MEMBER WADDY: But this measure is largely for documentation and not for -- I mean, I think it depends on what the measure is for.

CO-CHAIR TIRSCHWELL: I think all care measures, these process measures, are documentation burdens at some level because the process that they are talking about is an important one for standard of care.

MEMBER WADDY: I just think that, if it is for documentation, then I would see that; I would see it in a different light than if it is for improving quality. I mean, those are two totally different things, and you would use that same information in two totally different ways in this measure.

CO-CHAIR TIRSCHWELL: I think at high-quality centers it may be more of a documentation issue, and at lower-quality centers it may be really a quality issue.

MEMBER J. BAUTISTA: So, a minor point,

but there are some different exclusions, I think, in The Joint Commission measure compared to this measure. I think The Joint Commission measure excluded patients in clinical trials, that whole length-of-stay issue, and then patients admitted for elective vascular procedures, which was 10 percent of their population in The Joint Commission metrics. So, those are some differences.

CO-CHAIR TIRSCHWELL: A couple mor comments. Gail?

MEMBER COONEY: I just wanted to speak to the documentation versus improvement. Because I see this as much more focused on improvement than on documentation, because the numbers are so low that are being treated. I think the focus of this is that those eligible patients should be treated.

CO-CHAIR TIRSCHWELL: Daniel?

MEMBER LABOVITZ: Yes, this is perhaps a question for the developers. And I have got to say this discussion has gotten me smarter

than I used to be. So, I appreciate it.

This has to do with the diagnosis of ischemic stroke. We often, a few times a year anyway, see patients who had an ischemic stroke; a smart stroke neurologist would say, "Yes, that was a stroke," but the ED physician doesn't recognize it.

When is the diagnosis the diagnosis? Is it at discharge? Is it in those first two hours, and if you miss the diagnosis, you don't get a ding? How does that work?

CO-CHAIR TIRSCHWELL: I think it is at discharge. So, it is the final diagnosis. So, that would count against you if you get a lot of subtle posterior circulation strokes that your ED docs and even some of your neurology residents might miss on occasion. So, yes, you don't want to practice at that medical center.

MEMBER SUKO: I guess two comments, responding to that. From a practical perspective, the way it works in our facility

is it is at discharge, because it is abstracted by ICD-9 codes, which is what your coders code. And so, from a practical perspective at a facility, it really is at discharge.

And then, putting on my administration hat around quality, you know, there is the old adage about, if it is documented, it is not done. And the best way that we have to understand what you are thinking of as physicians as you are going through care is what you have documented.

And so, if there were a way that we could read your mind, other than your documentation, we could make those quality judgments. But, without that, we can't do that.

CO-CHAIR TIRSCHWELL: Okay. I am going to suggest we vote on impact.

(Laughter.)

I think our discussion has extended beyond impact.

Daniel, do you have another comment or you just left your name tag up? Thank you.

Okay. So, high impact?

(Vote taken.)

MS. THEBERGE: Fourteen high, 6 moderate, 2 low.

CO-CHAIR TIRSCHWELL: Okay. Moving on to 1c, evidence, Bill. This is still your measure.

MEMBER BARSAN: Yes, I think that we, again, six voting yes, one voting no on this. I think there was pretty good agreement with that.

CO-CHAIR TIRSCHWELL: The use of t-PA, that is measured here. We are voting.

(Vote taken.)

MS. THEBERGE: All right. Twenty responses. The clock is frozen. There we go, yes.

Eighteen yes, 4 no.

CO-CHAIR TIRSCHWELL: Okay. And then, moving on to 1b, performance gap.

MEMBER BARSAN: I think there is pretty good evidence that at least most people felt there was evidence that there was a performance gap here.

CO-CHAIR TIRSCHWELL: Okay.

(Vote taken.)

MS. THEBERGE: Twenty-one high, 1 moderate.

CO-CHAIR TIRSCHWELL: So, next is any comments on reliability. Bill?

MEMBER BARSAN: No, I don't think so.

MEMBER WADDY: Actually, can I make a comment regarding 1c really quickly for the evidence?

So, I just want it to be clear for future discussions and everything, this isn't evidence of utilization of t-PA. It is really what is the evidence for the information in this measure, right, which is different?

CO-CHAIR TIRSCHWELL: So, I think the evidence needs to be connected to the measure, but you don't have to prove that your measure

is improving outcomes. So, the evidence they give I think is mostly related to the benefit of t-PA, and you are asking whether that is appropriate or not?

MEMBER WADDY: I am really asking, the reason why I didn't put high for that one is because there are theoretical issues regarding the increased bleeding risks and the inclusion of the warnings and conditions. And so, that is why I didn't think it was as high. But if they had not put that section, then, of course, I mean, there is evidence for t-PA.

CO-CHAIR TIRSCHWELL: Yes, and I guess I would ask that NQF could qualify which area of evidence are we looking for here.

DR. WINKLER: When we are looking at the section under evidence, we are talking about that process of care. So, you are talking about in this particular case the use of t-PA. So, that is the evidence we are going to be judging the measure against. Once you get to scientific acceptability, you look at how the

measure is constructed and say, does it match that evidence that we know exists?

But the evidence is around the process of care that you are measuring. And we have got several measures that have tried to measure that process of care. The evidence should have been pretty much the same for all of them. It is how the measure works to reflect that.

MEMBER BARSAN: Yes. So, if we had issues with some of these things like we talked about with how these things are used and the warnings are listed, or whatever, would that come under this category? Any concerns about that? Yes? Okay.

DR. WINKLER: Yes, because you are talking about the measure specifications --

MEMBER BARSAN: Okay.

DR. WINKLER: -- under scientific acceptability of this measure's properties.

CO-CHAIR TIRSCHWELL: So, specifically, we are about to vote on reliability. You can

see up there it says precise specifications. I think the inclusion and exclusion criteria certainly are part of the specifications of the measure. So, that might be where you --

MEMBER BARSAN: So, I guess this is where people had issues, right.

CO-CHAIR TIRSCHWELL: Right. This might be where you want to put that emotion.

MEMBER GIDWANI: Perhaps I misunderstood, but I would assume that if there were issues with the denominator exclusions, that would be a validity issue rather than reliability. Okay. So, this is more about if different people assess the same dataset, do they come to the same conclusion? The validity would be the part at which we decide whether the denominator exclusions are appropriate.

DR. BURSTIN: Right, other than I think the fact that precision of specifications is still considered under reliability. So, if you don't believe some of the way it is even

written out as precise enough -- some of that will get picked up, I think, under validity. But I think it is a bit of splitting hairs.

MEMBER GIDWANI: Okay.

DR. WINKLER: One question on reliability?

CO-CHAIR TIRSCHWELL: Yes.

DR. WINKLER: In terms of the testing information that is presented, this is a facility-level measure, but the testing information says that it comes from four practice sites, large-group practice, a hospital-based neurology practice, a small-sized group practice, large-group practice.

So, I am having trouble understanding if we have tested the measure at the level of specification at the facility. And I would ask for some clarification.

DR. KATZAN: Yes, this was tested at a facility level. It was aggregated. Due to sample-size limitations for the most part, we don't test across physicians. So,

specifically, it would be aggregated across all facilities.

CO-CHAIR TIRSCHWELL: I think what Reva was asking, if it is a group practice, that doesn't sound like a facility. I don't know. Is that --

DR. WINKLER: That is what I was wondering.

DR. KATZAN: I mean, all I will say is we aggregated across facilities. We don't do by individual physicians.

DR. WINKLER: I guess if this measure is meant to measure at the level of the facility, the typical facility is a hospital. So, how does this measure give you results of one hospital that you would compare to other hospitals, and yet, when you have tested it and used it in practices? That is where I am confused.

CO-CHAIR TIRSCHWELL: Any further comments from the developer? So, I guess we are suggesting that reliability data that is

presented doesn't quite match up with the facility-level testing that is proposed.

MS. YODICE: Well, I mean, I will say that there was 100 percent agreement if you look at the sample size of 148 in our inter-rater reliability.

CO-CHAIR KNOWLTON: I think the question that is being put is that this is being described as a facility measure.

MS. YODICE: Right.

CO-CHAIR KNOWLTON: So, then, you would have to standardize your reliability on a facility level. And if I understand the data correctly, you looked at reliability, inter-rater reliability, and so forth, at a physician level and/or at a physician-practice level, not at a facility level, which is usually a hospital. Is that correct? Or do we misunderstand? I think that is the question.

MS. YODICE: Right. In this case, it shouldn't necessarily make a difference

because --

CO-CHAIR KNOWLTON: You can make that argument, certainly. But answer the question first. Is there facility-level hospital data in this measure as part of the reliability?

MS. YODICE: I don't know the answer to that question.

CO-CHAIR KNOWLTON: You don't know? You can go on and make your point. I didn't mean to interrupt you, but I wasn't getting an answer to the question.

MS. YODICE: Okay. So, what I do know is that the measures technically, they are tested at a facility level based on random selection from the facility.

CO-CHAIR KNOWLTON: What is facility level? You mean like a practice?

MS. YODICE: Yes. Like a site, a test site, a hospital.

CO-CHAIR KNOWLTON: It is out of a hospital?

MS. YODICE: Yes.

MEMBER WADDY: But this says it was based on practices. I mean, in some places a practice may be at a hospital, but in other places, such as Fairfax, there are like three different practices within a hospital.

MS. YODICE: Okay. Yes. These are inpatient measures only.

MEMBER WADDY: Right, but there may be three different practices, three different physician practices within a hospital. And so, those may be completely different, such as the INOVA Group or the Richard Seestedt group at INOVA may be doing one thing; whereas, the Eric Sklar Group at INOVA does something completely different.

CO-CHAIR TIRSCHWELL: It seems like there is ambiguity about the reliability as presented. I don't know; do you have anything further to add?

MS. YODICE: Well, as I was saying before, I know these are inpatient measures only. So, they were tested across four sites,

technically tested at facility level based on random selection from the facility. We don't test across physicians in this testing model due to sample-size limitations.

So, essentially, I mean, that is all I can say to answer the question.

CO-CHAIR TIRSCHWELL: So, I guess we have to individually decide whether these data referring to practices are reliable enough a measure of facility-level reliability and vote according to our conscience. I am not sure there is anything else to add to that.

So, could we go ahead and vote on reliability?

(Vote taken.)

MS. THEBERGE: There we go.

One high, 12 moderate, 4 low, 5 insufficient evidence.

CO-CHAIR TIRSCHWELL: So, we continue.

MEMBER BARSAN: Yes, onto validity, I don't think there were any major issues raised in the group about this specifically. We had

4 high, 2 medium, 1 low votes in terms of the validity.

CO-CHAIR TIRSCHWELL: And again, you know, the issue about the specifications of the exclusions may be relevant in your vote at this juncture.

Any other comments before we move to vote?

MEMBER KAPINOS: So, I am confused. The overall assessment was that, actually, a lot of us did not like the fact that it was so with all those exclusions or contraindications and the other term that they used.

So, here, for 2b.1, specifications are consistent with the evidence, I should still say yes because that is exactly what is written in the guidelines, right? But, actually, we are thinking it is too specific and we prefer, actually, the other. The Joint Commission submitted one that just leaves it up to the clinician to exclude the patients just by documenting, and they did not list all

those contraindications to t-PA.

So, here I am going to score even if in my heart at the end I am going to maybe vote no because I prefer the other one that does leave more to the clinicians' judgment? Or am I a little bit confused on it.

CO-CHAIR TIRSCHWELL: Yes, and I think there is some ambiguity in the process a little bit. But let me just comment quickly that we heard a lot of emotion about the specifications, and we have not really seen that reflected in any of the voting just yet. So, I think that should be part of your voting along the way. And if at the end you are going to vote no, I would hope that at some point along the way you voted down one of the criteria. I think that would be consistent.

At this point, we should take each measure independently, not worry about whether we like The Joint Commission one better or not, and just kind of take this one on its own merit.

DR. BURSTIN: Just a thought on that, David.

CO-CHAIR TIRSCHWELL: Go ahead, yes.

DR. BURSTIN: So, just to make that even more clear perhaps, vote each individual measure as they stand based on the criteria before you. At the end of the day, when we have that session outlined for competing and related measures, we will actually show you your votes across the competing and related measures, so you can actually see how you rated them on each of the criteria and overall to help make that assessment. But, for now, just vote on the individual measure.

CO-CHAIR TIRSCHWELL: Yes, sure. Go ahead.

DR. ANTMAN: I don't mean to put our Joint Commission colleagues on the spot, but I am just trying to understand why there are some questions about specific items that we have listed as exclusions. And as Diedre explained, we actually regard these as

discretionary exceptions from the measure.

I am a little unsure as to why there are questions about individual items listed there, given that we are fairly certain that we match The Joint Commission list.

If I may, either to Karen or Ann at The Joint Commission, is it perhaps that that full list of contraindications, et cetera, other possible exceptions to the measure, are those bundled under what has been listed in The Joint Commission measure as documented reason for not initiating IV t-PA?

CO-CHAIR TIRSCHWELL: Just as a point of process, we are probably going to need to go through the Committee and not directly talk to each other as side participants. But go ahead and answer that question, if you will.

MS. KOLBUSZ: Yes, I think this came up on the Work Group conference call as well. We have a data element reason for not initiating IV thrombolytic therapy. The guidelines for instruction are included in the data element

definition for that particular data element.

And the list, the table which is called "Conditions making the administration of IV thrombolytic therapy inadvisable" is reference material provided in our measure information form. Those are not necessarily absolute exclusions to the measure. We still require that, for any reason for not initiating IV thrombolytic therapy, that it be explicitly documented in the medical record by the physician, the APM, the physician assistant, or pharmacist.

Therefore, if it is in the table, which was taken directly from the FDA list of contraindications, from their labeling instructions, or if they had another reason for not giving t-PA, that that would all roll into that data-element reason for not initiating IV thrombolytic therapy.

CO-CHAIR TIRSCHWELL: Okay. So, just to clarify, it sounds to me like the lists are likely identical, if we got them all out and

put them next to each other. And I think that is your point, Mark. But the reality is these guys put it on the form that we looked at in more detail and are perhaps inappropriately suffering as a consequence of that. And so, we may want to consider whether that is appropriate or not. My guess is that these criteria on a list are the same, no matter what.

MEMBER KAPLITT: Can the developer explain to me the answer to 2b.3, that at the time of testing the measure did not have exceptions? Page 17. Because we have been talking about all these exceptions, and we are talking about the testing that was done in these various settings of unclear relevance. And it explicitly states earlier, I think, in -- I forget -- 2b.1 or something, that they are supposed to collect exception data which I don't see presented. So, I don't understand that answer.

DR. ANTMAN: Right. So, I recognize

that that is confusing.

So, the measure that is under review currently, the T-PA-administered measure, this is a new measure for us in that the pair that we presented, the measure previously looked at was closer, if you will, to the measure that is now designated as the initiated measure. Because as our Work Group discussed it, as we were updating the measure set, the feeling was that we should separate out the consideration from the initiation.

So, the testing that was done on these measures was on the previous version that was, in fact, the considered measure, but the testing was done at the data-element level that we felt was still relevant to this new measure, the initiation measure.

MEMBER KAPLITT: Well, just to clarify, because I am a neurosurgeon, so I think very simply. Okay? So, what you tested is not what we are being asked to approve here today?

DR. ANTMAN: So, what we tested is not

an exact match to this measure, that is correct. We tested at the data-element level, but we felt that the data elements matched this measure as revised.

MEMBER KAPLITT: But exceptions were not part of that?

DR. ANTMAN: That is correct.

MEMBER KAPLITT: Which is the subject under discussion for the last 20 minutes, as I understand it, right?

CO-CHAIR TIRSCHWELL: And so, in fact, to your question, Michael, these are inaccurate answers at this point, that there are no exclusions. Is that also right, Dr. Antman?

MEMBER KAPLITT: That is the point. So, the point I am trying to clarify is -- and this relates to the vote we even just took, in my view -- but the data we are being asked to evaluate tested a measure that did not include exceptions. The measure we are being asked to approve does include exceptions, and that has

been the major subject of discussion.

So, what I am questioning is, how can we possibly evaluate that data if it is not actually including all the elements of the measure we are being asked to evaluate here, unless I am misunderstanding this answer? And it doesn't sound like I am.

DR. ANTMAN: No, I don't think you are misunderstanding. Again, it was tested in the earlier version. We felt the testing was valid for this measure because it was tested at the data-element level. But it is correct that in the version that it was tested it did not have the exclusions or the exceptions.

CO-CHAIR KNOWLTON: I do have a comment. The developer has a burden, I think, of bringing the evidence, the assumption they made that these two things are comparable, there is a burden to show that, to carry the question of validity, which I think is your point. I agree.

CO-CHAIR TIRSCHWELL: Any other comments

related to validity issues before we go ahead and vote?

MEMBER KAPINOS: I just wanted to make the comment that I told you, to share it with everybody. The concern that I have with all this list clearly-stated in that measure is that, then, it is going to be easy for a physician in the ER to think about, "I want to make my hospital have a good score. Therefore, I am going to use each of these things to exclude as many patients as possible. Therefore, my hospital will have, yes, a very good measure, but I am going to give less t-PA, actually."

Because here it is clearly stated which criteria I can use to exclude my patients from receiving t-PA, as opposed to if you make it more blurry, like documented evidence against t-PA, like on the other measure, the one proposed by The Joint Commission. Actually, you are not giving a clear checklist to a physician to be used to exclude the patient.

So, I mean, I think it is less facilitating, malevolent people that would actually exclude more patients with the interest to make the hospital measure look better.

CO-CHAIR TIRSCHWELL: You know, I guess that is a possibility. I have a hard time believing that emergency medicine physicians would really stack the deck like that.

I mean, vote your conscience. If you really think that is a big issue, let it be known. But my personal opinion is they have got a checklist already. Almost every place does. And it is not that hard to find a reason to exclude someone. I don't think this is going to really make the difference.

Salina?

MEMBER WADDY: Actually, I mean, I think you are right regarding the physician and the people who are the treatment teams, but this is really a facility-level indicator. And so, whether or not that changes things or not, I think potentially it could.

MEMBER KAPLITT: Fundamentally, what we are doing here is speculating, right? I mean, we are supposed to be basing it on data. And my point is we can all have our various opinions of how this could be used or misused, but the problem that I see here is that we don't have the data. If I have a bias that this could be misused, and the data says that that is not happening, fine, that I would have no problem with that. But we don't have the data. That is my concern.

CO-CHAIR TIRSCHWELL: Yes, go ahead, Gwen.

MEMBER BUHR: So, The Joint Commission measure has a separate document where it outlines all of the exclusions, where if you are going to follow the measure, you have this separate document. And it has all of the same exclusions as this measure exactly. So, I just wanted to make that point.

CO-CHAIR TIRSCHWELL: I think that is a good point, but they also have data showing

that it is --

MEMBER BUHR: Yes, exactly.

CO-CHAIR TIRSCHWELL: -- reliable and valid in its --

MEMBER BUHR: So, the issue is that this measure doesn't have the validity data exactly as presented. I don't know if it really matters, the different -- I mean, they are the exact same things, but they haven't presented us with the data.

CO-CHAIR TIRSCHWELL: Any other comments?

Gail is ready to vote.

I will give the developer one last chance to respond. Then, I think we need to vote.

MS. HANLEY: Yes, just one thing to add. As Mark explained, we tested the prior measure, the t-PA considered measure that wasn't previously endorsed, at the data-element level. While that measure did not have exceptions, information was collected and

tested on the reasons why a patient may not have received t-PA.

So, in the prior measure, the considered measure, you were counted in the numerator if you either received it or if you didn't receive it, but, yet, there was documentation for those reasons. And in our testing table, we actually have some examples which look quite similar to the list of the exceptions that we included in this measure submission form to reasons why the patient was not prescribed t-PA.

So, I just thought I would add that. Even though the data elements weren't tested as exceptions in the prior measure, the data elements were collected and tested.

CO-CHAIR TIRSCHWELL: Is that data on the forms?

MS. HANLEY: I would have to look closer. I don't think we actually included that.

MS. YODICE: This is the actual tool

that the abstractors used when they were abstracting the data elements from the medical records. So, it is the list of inclusions.

MS. HANLEY: But we didn't provide that.

MS. YODICE: It is not in the form, no.

CO-CHAIR TIRSCHWELL: I think we are just going to have to vote on what we have seen and what we have discussed.

2b, validity, I think we should open for voting.

(Vote taken.)

MS. THEBERGE: One high, 5 moderate, 5 low, and 11 insufficient evidence.

CO-CHAIR TIRSCHWELL: So, okay, I think that reflects the discussion pretty well. So, it suggests that -- I think is this a full stop at this point? Okay. So, this measure did not meet criteria to move forward, either.

So, at this point, I think we should take a break, although we are way behind schedule. So, if we could try to limit it to a five-minute break, I think that would be

great, and we will retool about how much we are going to try to get done before lunch.

(Whereupon, the foregoing matter went off the record at 11:49 a.m. and resumed at 11:58 a.m.)

CO-CHAIR KNOWLTON: Okay. Welcome to 1952.

(Laughter.)

Dan? Where's Dan? He is over here.

MEMBER LABOVITZ: So, Measure 1952 is a new process measure proposed by the American Heart Association, again focusing on delivery of t-PA, but this time not "Did you give it or not or did you think about it?", but if you gave it, how fast did you deliver? Basically, three time points needed: time patient was last-known well, time patient arrived in the hospital, and time t-PA was actually started.

There have been several publications since the original t-PA trial showing that, although originally we didn't recognize it, time does matter. Most recently published in

circulation was the review of more than 25,000 patients who were given t-PA with their data incorporated into the Get With The Guidelines database.

Looking there, the faster the delivery, the better the chance of a good outcome and the lower the risk of mortality. Other meta-analyses on trials using t-PA beyond the three-hour window have also shown a real time effect.

The Work Group thought that this was an important measure unanimously for impact on 1a. We had no concerns about that.

CO-CHAIR KNOWLTON: Any questions on impact for Dan or the Committee or thoughts?

(No response.)

Okay. Let's go to a vote.

(Vote taken.)

MS. THEBERGE: We are at 19. Twenty-one.

Twenty-two high.

CO-CHAIR KNOWLTON: Okay. Unanimous.

Go ahead, Dan.

MEMBER LABOVITZ: For 1c, the rationale, again, the same data, the same concept.

It was broadly agreed that shorter is better. However, there are concerns on this point about whether -- and I am not sure it fits into this area -- but concerns that a rush to treat might lead to poor treatment decisions. That was a subject of much discussion within the Work Group.

Also, some concern about centers that may do a more sophisticated evaluation of patients that might deliver true quality, might deliver better care, and cost a little bit of extra time.

On the whole, though, the group agreed that time trumps everything else. And it was unanimous in favor.

CO-CHAIR KNOWLTON: Comments or questions for Dan? Or comments to the group?

(No response.)

Okay. Let's vote.

(Vote taken.)

MS. THEBERGE: We are at 20. We still need one.

Twenty-one yes, 1 no.

CO-CHAIR KNOWLTON: Okay. Dan?

MEMBER LABOVITZ: Okay. For 2a, the scientific acceptability of the measure --

CO-CHAIR KNOWLTON: 1b.

MEMBER LABOVITZ: Oh, I'm sorry. Have I skipped 1b? I do apologize.

Performance gap, opportunity for improvement. The American Heart Association has published data showing that nationally in the group of hospitals that is reporting data the delivery of t-PA within 60 minutes is at an abysmal 25 percent. The Target Stroke Initiative by the American Heart Association is pushing to raise that number to at least 50 percent. So, there is a wide gap there in terms of delivery in a timely fashion.

CO-CHAIR KNOWLTON: Comments? Questions?

(No response.)

Vote?

Wait for her to get the clock up there. There you go.

(Vote taken.)

MS. THEBERGE: We have 19 responses.

Twenty-one high, 1 moderate.

CO-CHAIR KNOWLTON: Okay. Evidence. Oh, we did evidence. We are in reverse. I'm sorry.

MEMBER LABOVITZ: So, I think we are at scientific acceptability of measure properties.

CO-CHAIR KNOWLTON: Right.

MEMBER LABOVITZ: There was again discussion here about whether including the 3-to-4.5-hour stretch was scientifically-valid. It is not something that has been endorsed by the FDA.

The group, though, agreed that this is a measure of how quickly you deliver t-PA. Whether your hospital or institution or you

personally think you should be delivering t-PA beyond three hours is not measured by this measure, and there is no ding or quality assessment on that point.

There were no other comments on that area.

CO-CHAIR KNOWLTON: Can I ask you a question on that, then?

MEMBER LABOVITZ: Yes.

CO-CHAIR KNOWLTON: I read in the scientific acceptability area the issue of risk stratification and whether that was pertinent to t-PA. I don't know the answer, but I wonder if it was answered in the group.

MEMBER LABOVITZ: I am very happy to bloviate on that point.

(Laughter.)

But I am not sure that -- I think that is, again, a discussion that is maybe served over a bottle of wine.

When is it that you should stratify hospitals?

CO-CHAIR KNOWLTON: I am okay with that, by the way.

(Laughter.)

MEMBER LABOVITZ: Oh, okay.

I can certainly say I work in a medical center that has three separate institutions. One of them, many more patients arrive early than at another. The patients behave differently. It is easier to deliver t-PA at one institution, not because the doctors are different or the institutional practice is different, but the patients are different.

I still think that stratification on any of these points risks losing the main focus. And that is time is of value.

I think that delivering t-PA within one hour is a goal that is reasonable and achievable, no matter where you are.

MEMBER KAPLITT: I would note that there are some reasonable exceptions that are included in the measure for delay. So, you know, if something happens, it won't count

against you if it is reasonable.

MEMBER LABOVITZ: Yes, there are several reasons why you might exclude a patient from this measure. Even though t-PA was delivered, it was delivered late, but it was delivered late because you had to call Mexico to find out that the patient didn't have an intracerebral hemorrhage. That certainly happens to us in The Bronx.

CO-CHAIR KNOWLTON: Jocelyn?

MEMBER J. BAUTISTA: I just want to clarify. Are we actually talking about two different measures here? Because they talk about the percentage of patients who receive t-PA within 60 minutes, but, then, there is also this second sentence of median time from hospital arrive to t-PA administration. So, are these two separate measures that we are evaluating?

MEMBER LABOVITZ: That might be a better question for the developers. I can comment just to say I think the data acquired is

exactly the same. It may really just be two different ways of reporting it.

DR. SCHWAMM: Yes, so speaking for the measure developer, the measure is the percentage of patients who get it within 60 minutes. But, in addition, the median time to treatment is reported for centers to make use of that information to understand where they are in relation to their time-to-treatment goals.

The other comment that I would make is that, during the conversation, there was a request by the Working Group that we provide the other NQF-endorsed measures that have a similar set of time requirements for thrombolytic therapy. We have those available for you, if you would like to take them in writing.

But it is the primary PCI within 90 minutes of hospital arrival, which is NQF 0163; thrombolytic therapy within 30 minutes of arrival, which is NQF 0164, and time to IV

thrombolytic therapy, all for acute MI. NQF pending 1952 is this measure.

You can see that the measure exclusions are all the same. They were modeled specifically after the NQF endorsement approach of listing reasons that relate to patient factors. So, I just wanted to follow up on that.

CO-CHAIR KNOWLTON: Jocelyn, did that answer your question?

MEMBER J. BAUTISTA: Yes.

CO-CHAIR KNOWLTON: Anybody else?

(No response.)

Okay. Ready to vote on reliability.

(Vote taken.)

MS. THEBERGE: Twenty-one responses. Okay.

We have 14 high and 8 moderate.

CO-CHAIR KNOWLTON: Okay. Dan?

MEMBER LABOVITZ: Okay. I think that takes us along to No. 3, usability.

CO-CHAIR KNOWLTON: Validity.

MEMBER LABOVITZ: 2b, validity. I have skipped a piece.

We had high ratings on this point; no specific comments.

CO-CHAIR KNOWLTON: Questions or comments from the Committee?

(No response.)

Ready to vote.

(Vote taken.)

MS. THEBERGE: We are at 20. We are at 21.

Twenty high, 2 moderate.

CO-CHAIR KNOWLTON: Okay. Dan? Usability?

MEMBER LABOVITZ: Okay. Usability, again, the Work Group thought that this was universally high. One voted medium.

The only comment there was that this seemed like a new measure, so maybe we don't know enough about it. I do think that 25,000 t-PA cases is a large number. I think we do know a lot about it. But that is just me.

Overall, the group thought that this was highly-usable.

CO-CHAIR KNOWLTON: Comments? Questions?

(No response.)

Okay. Let's vote.

(Vote taken.)

MS. THEBERGE: We are still at 21 responses.

Twenty-two high.

CO-CHAIR KNOWLTON: Feasibility?

MEMBER LABOVITZ: Feasibility. Some concern about the fact that not all of these data are routinely collected, but at least for centers using Get With The Guidelines they are routinely collected.

I can certainly say, as a data guy, time stuff is hard. It is easy to screw it up. But we are doing time all the time. So, I don't think there is an escape.

The group thought that this had high feasibility with one vote for medium.

CO-CHAIR KNOWLTON: Questions or comments?

(No response.)

Ready to vote.

(Vote taken.)

MS. THEBERGE: We are at 21. We are still at 21. Can everyone re-vote? Has anyone decided not to vote? We are still at 21.

CO-CHAIR KNOWLTON: You will be punished.

(Laughter.)

MS. THEBERGE: All right. We have 16 high, 5 moderate, and 1 low. That's 22.

MEMBER KAPINOS: Is there any ongoing like attempts to standardize the way we report time? Because, as we know, we have the time in the computer, on the video when we monitor patients for seizure, the time on your EKG machine. All those times are different with many hours sometimes at some institutions between each modality that you are using for

the time.

So, actually, with that measure in place, hopefully, in the future that will also foster some initiatives so that hospitals standardize and align their times in all the different machines that we use. Because for the documentation and the work of the physician, I mean, it is a mess.

CO-CHAIR KNOWLTON: You may want to have that discussion with Dan over the bottle of wine he promised.

(Laughter.)

MEMBER LABOVITZ: I am buying tonight. Everybody is invited, yes.

(Laughter.)

CO-CHAIR KNOWLTON: So, the overall. That one passed; we go on to the last issue, right? We haven't changed the screen yet. We did feasibility.

So, we are on to overall suitability for endorsement.

MEMBER LABOVITZ: On overall

suitability, the Work Group voted unanimously after discussion. We had to twist one arm, but it worked out.

CO-CHAIR KNOWLTON: Questions or comments?

(No response.)

Okay. Let's vote.

(Vote taken.)

MS. THEBERGE: Twenty-two yes.

CO-CHAIR TIRSCHWELL: All right. The next measure for review is 0438. That is a Joint Commission measure, antithrombotic therapy by the end of hospital day two.

Greg, do you want to introduce that?

MEMBER KAPINOS: So, actually, it is only afterwards that I came up with a question. I thought it was a straightforward definition. So, all these patients should receive antithrombotic therapy within the first 48 hours, and that is the definition that is used for that measure.

However, I couldn't find out the

definition of which antithrombotic regimen are we talking about because I am guessing that everybody is talking about like ischemic, I mean secondary -- it is preventing the clot from progressing and initiating early secondary prevention of ischemic shock with an antiplatelet agent or an anticoagulant if you have some other indication for anticoagulation rather than antiplatelets.

But the fact that it is a broad terminology, I just have a question to clarify that before we go on to the votes. Because it would be inappropriate to count, also, as antithrombotic, for instance, VT prophylaxis. So, if I put my patient on hep subQ at day one, I am going to match that measurement, I believe, even if I did not give aspirin for the stroke.

So, is it detailed somewhere? I missed it? Okay.

MEMBER J. BAUTISTA: I have the data dictionary, if you -- it excludes subQ

heparin. So, that is not considered.

MEMBER KAPINOS: Okay. And so, why didn't we define better like aspirin, like in the guidelines they tell you the use of aspirin between 150 milligrams and 320 milligrams is exactly what we are looking for for the ischemic strokes or coumadin -- well, not coumadin within two days. So, actually, that is the only regimen.

So, why didn't we replace antithrombotic therapy by aspirin 150 milligrams up to 320?

CO-CHAIR TIRSCHWELL: I think, on a process point, the additional materials that are submitted with these measures are part of their submission. And they do go into detail.

Do you have it in front of you? Can you read out all of them, Jocelyn?

MEMBER J. BAUTISTA: Well, I have the data dictionary in front of me that they refer to in the submission. They simply say antithrombotic therapy. They don't list antithrombotic therapies, but they say

exclusions from antithrombotic therapies, heparin subQ, heparin flush, and Hep-Lock.

CO-CHAIR TIRSCHWELL: It must have been another measure that I saw a specific list for.

MEMBER J. BAUTISTA: Oh, you know what? There is. There is an Appendix C, Table 8.2.

CO-CHAIR TIRSCHWELL: You know, I think they don't specify that it is just aspirin. They really include all of the sort of full-dose antithrombotics. And it sounds like they call out that DVT prophylaxis dosing is not an appropriate measure to meet the exclusion.

Do the developers have anything to add to that?

DR. SCHWAMM: Yes, I would emphasize just what you have said, that all FDA-approved or guideline-endorsed antithrombotic agents are listed in the coding instructions and the data dictionary that is available to abstractors.

But because there is confusion out there

around the difference between antiprophylaxis for DVT, which is not appropriate antithrombotic therapy -- and quite frankly, the reverse, that an aspirin a day does not prevent DVT. Those are called out specifically, so that there is a high reliability in the implementation of the measure.

CO-CHAIR TIRSCHWELL: So, theoretically, that is in place, although it was buried a couple of layers down.

MEMBER KAPINOS: Let's look at 1a?

CO-CHAIR TIRSCHWELL: Impact.

MEMBER KAPINOS: Yes, let's vote for impact. There was no specific question.

CO-CHAIR TIRSCHWELL: Go ahead and open the voting. Okay.

(Vote taken.)

MS. THEBERGE: Twenty-two high.

CO-CHAIR TIRSCHWELL: Great. Go on to 1c.

MEMBER KAPINOS: Okay. For the evidence

supporting this, also, everybody was compelled that it was high evidence. No specific question.

CO-CHAIR TIRSCHWELL: Let's vote on 1c then.

(Vote taken.)

MS. THEBERGE: We are still at 21. Twenty-two high -- I'm sorry -- yes.

MEMBER KAPINOS: We have a question.

CO-CHAIR TIRSCHWELL: And then, next is --

MEMBER KAPINOS: Hold on. We have a question from Risha.

CO-CHAIR TIRSCHWELL: Yes, please. Go ahead, Risha.

MEMBER GIDWANI: Thanks.

This is actually a comment rather than a question. I notice here, and in many other places in the measures, when grading the quality of the evidence, folks are using guidelines indicating that the highest level is that from multiple randomized control

trials or meta-analyses.

And for the developers, I would suggest actually disaggregating those two categories. Meta-analysis is not necessarily of the same level of evidence as multiple randomized controlled trials. The quality of a meta-analysis is usually akin to that of an observational study.

Depending on the quality of the meta-analysis, the randomized controlled trial can actually trump that in terms of data validity. I will refer folks to a Statistics in Medicine article from Shuster in 2010.

I am happy to discuss this offline with any developers, but just a comment and suggestion to be made.

CO-CHAIR TIRSCHWELL: I guess I would only add that meta-analyses can be of randomized trials or observational studies, and they should specify that.

MEMBER GIDWANI: I agree with that. Even when the meta-analysis is, though, only

using randomized controlled trials, it is considered to be, the meta-analysis itself of the RCTs is considered to have the quality of evidence of an observational study. So, suggestion to disaggregate those two levels.

CO-CHAIR TIRSCHWELL: Okay. So, that was a comment on the previous measure as to a performance gap, Greg. Any comments?

MEMBER KAPINOS: Oh, of course. Sorry. For 1b, yes. For the performance gap, yes, people were considering the fact that now that actually we matched -- 97 or almost 100 percent of the hospitals are with good compliance with this measure. Actually, people were saying the 1b item will be ranked very low because you probably cannot improve from 98 to 100 percent.

However, then, I think the Work Group suggested that, even a minimal increase from 99.5 to 99.6 increment, since we are talking about such a powerful impact from initiating that therapy, will link to better outcomes.

Actually, it is still important to have this measure ongoing and maintained. So, overall, I think everybody was willing to give it a high score.

CO-CHAIR TIRSCHWELL: Risha, do you have another comment?

MEMBER GIDWANI: No.

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

MEMBER J. BAUTISTA: No.

CO-CHAIR TIRSCHWELL: Put your name tag down, please, then.

Yes, Ramon?

MEMBER R. BAUTISTA: But the reason for the gap was not given, though. In other words, you could have a 2-percent performance gap, but that wasn't even stated, right? So, it is only known, but the reason for that was not stated, either.

CO-CHAIR TIRSCHWELL: It shouldn't be an explanation for your performance gap that would exclude you from the measure, right.

So, I am not sure that is --

MEMBER R. BAUTISTA: Just to make sure, what that actually taken into account, though, in this study? Did they exclude people who could not take --

CO-CHAIR TIRSCHWELL: Documented reason for not administering anti-thrombotic therapy by end of hospital day two. So, yes.

You know, the other thing that I would add that was brought out in the application was that the Get With The Guidelines hospitals are the hospitals which have self-selected and committed to high-quality stroke care. So, even if the Get With The Guidelines hospitals are performing at a very high level, you can bet -- and again, I am making an assumption -- but you can bet that the hospitals that aren't participating may be performing at a lower level. And having this endorsed as an important measure is probably even more important than perhaps some of the data that we are not able to observe at this point in

time.

And then, I guess I would ask the NQF people, if we all agreed that it had topped-out and we wanted to put it on reserve, when would we say that, at the end after it is approved?

DR. BURSTIN: Soon.

(Laughter.)

We are getting there.

But just a question for The Joint Commission. It looks like you actually give two rates, the first of which is based on Joint Commission data, not Get With The Guidelines, and that is still at 96-97 percent. Is that correct? I just want to clarify it.

So, actually, they are both high, even outside the Get With The Guidelines.

CO-CHAIR TIRSCHWELL: But The Joint Commission data are only in primary stroke centers. It might even be more selective than Get With The Guidelines.

DR. SCHWAMM: Right. Just to clarify that, there are 1600 hospitals in Get With The Guidelines currently. Eight hundred of those are primary stroke centers.

So, I think what you are seeing is a sort of gradient effect there where we would expect to see the highest performance among primary stroke centers, but we are still only capturing a third of U.S. hospitals in the data presented.

DR. BURSTIN: So, essentially, just to answer David's question, we would ask you to finish all of the first criteria here, the three subsets. You need to go through evidence as well.

We would, then, ask you to then pull up a slide and see, do you think -- and actually, you would have to go through scientific acceptability as well. So, the measure has to be rated highly on scientific acceptability, and, otherwise, would meet importance to measure and report, with the exception of the

fact that the Committee thinks it has topped-out.

And then, you could consider moving that into reserve status. What that means is that it is technically still endorsed, but there is probably an expectation that continuing to push on that -- and again, the subset of hospitals issue is an important one -- may not be the best strategy.

CO-CHAIR TIRSCHWELL: So, if the performance gap was the only low score, that is when we would consider that situation?

DR. BURSTIN: Please finish the evaluation first.

CO-CHAIR TIRSCHWELL: Okay. So, let's go ahead.

Yes? Oh, sorry, Gail. Go ahead.

MEMBER COONEY: I just want to clarify with The Joint Commission. Of the data that you are reporting, what proportion of those hospitals are being evaluated as primary stroke centers? Do you have hospitals that

aren't primary stroke centers that are in that data or no?

MS. KOLBUSZ: Yes, we do give hospitals that are primary stroke centers in the data reported. The data came from our ORYX core database. So, there is approximately 170 hospitals in that data collection.

MEMBER COONEY: And what proportion of those hospitals are primary stroke centers?

MS. KOLBUSZ: About half.

DR. BURSTIN: Is there a difference in performance between those that are and those that aren't, I guess is the obvious next question. You set it up.

MS. WATT: You know what? It is an analysis that we haven't done, but it is a good one. And thank you for suggesting it.

CO-CHAIR TIRSCHWELL: Daniel?

MEMBER LABOVITZ: This measure did give me some heartburn. Because I spend, actually, a lot of reviewer time gathering these data. We put a lot of effort into getting this data

electronically. And we have given up trying to fix our mistakes, which accounts for most of the 2 to 3 percent of patients where we are not treating with aspirin. It is a glitch in the system. It is a misreading of documentation. It is just a coding error.

I think that it is expensive to gather data that you are not using. And I think any center that is gathering these data is doing well. I also think, to speak to Dr. Trischwell's point about centers that may be doing a bad job, we don't have the data on them.

So, I think in previous discussions we have said, "Hey, let's not make it up here." We haven't been shown the data to suggest that this measure is doing anything. And, boy, we are topped-out.

I think somebody in our Work Group discussion said, "Where is the NQF Hall of Fame in endorsement, but not something that CMS picks up and leaves us with for the next

30 years?"

MEMBER KAPINOS: So, let's vote?

CO-CHAIR TIRSCHWELL: Thank you.

(Vote taken.)

MS. THEBERGE: Three high, 8 moderate, 9 low, 2 insufficient evidence.

CO-CHAIR TIRSCHWELL: So, it is a dead heat.

NQF, how should we proceed?

DR. WINKLER: Well, especially if your conversation is around potentially going to reserve status, it almost doesn't matter because either it passed and you go on or it didn't pass but you are contemplating reserve status and you would go on. Regardless, you are going on and finish the rest of the evaluation.

CO-CHAIR TIRSCHWELL: Okay, then, 2a, reliability.

MEMBER KAPINOS: No issues.

CO-CHAIR TIRSCHWELL: Comments or questions about reliability for this measure?

(No response.)

Let's go ahead and vote then.

MEMBER J. BAUTISTA: But wait. Didn't of the Work Group members write "Reliability testing data not clearly presented."?

MEMBER KAPINOS: Yes, I don't remember exactly where it was at. So, if you want to go ahead -- who put that?

CO-CHAIR TIRSCHWELL: I think that was me, quite honestly, and it has to do with, as was already commented on, the issue surrounding the way some of the data come out and the denominators. I will have to pull it up.

You know, quite honestly, I was satisfied after I got further information. Those comments were from sort of before the Work Group meeting. so, I don't know that we need to -- did anybody else have any question about that?

(No response.)

MEMBER KAPINOS: See also the comment

about the new anticoagulants, or are they in that longer list in the dictionary?

MEMBER J. BAUTISTA: Yes.

CO-CHAIR TIRSCHWELL: Yes.

MEMBER J. BAUTISTA: There are about 150.

MEMBER KAPINOS: Okay. That is why I didn't go over it. Okay.

CO-CHAIR TIRSCHWELL: So, let's go ahead vote then on reliability.

(Vote taken.)

MS. THEBERGE: We are at 20 responses. Sixteen high, 6 moderate.

MEMBER KAPINOS: Moving on to validity, I don't recall any specific concerns.

CO-CHAIR KNOWLTON: Okay. Any other comments?

(No response.)

Let's go ahead and vote.

(Vote taken.)

MS. THEBERGE: Twenty responses. Twenty-one. Can everyone vote one more time?

CO-CHAIR KNOWLTON: There we go.

MS. THEBERGE: There we go.

Sixteen high, 6 moderate.

CO-CHAIR KNOWLTON: So, moving along, usability.

MEMBER KAPINOS: The same thing, I don't recall any specific concerns. It is pretty easy to --

CO-CHAIR KNOWLTON: Any comments or questions about usability?

(No response.)

Let's go ahead and vote.

(Vote taken.)

MS. THEBERGE: Nineteen high, 3 moderate.

CO-CHAIR KNOWLTON: Okay, and feasibility? Any comments or questions?

MEMBER KAPINOS: No.

CO-CHAIR KNOWLTON: No?

Yes, please, vote.

(Vote taken.)

MS. THEBERGE: Twenty-one responses.

Eighteen high, 3 moderate, 1 low.

CO-CHAIR KNOWLTON: Okay. And then, I think just the overall suitability is the next vote. Any other comments before we vote on overall suitability?

(No response.)

I don't see any. Let's go ahead and vote.

(Vote taken.)

MS. THEBERGE: Twenty-one responses. Okay.

Nineteen yes, 3 no.

CO-CHAIR KNOWLTON: Okay. So, I guess now we have to have the conversation about reserve status. How do we start that conversation? Oh, here, criteria for reserve status.

So, evidence, is there strong, direct evidence to a desired health outcome? I think the answer to that is yes.

MEMBER KAPLITT: Can you explain reserve status for a second, so that we understand it

better?

DR. BURSTIN: So, reserve status would be reserved, I guess, for measures that otherwise would pass all the NQF evaluation criteria. They are evidence-based. They are important. They are scientifically-valid. But they are at such high levels of performance that the question is, does continuing to be measuring them on a regular basis, as opposed to more periodic surveillance and assessment to make sure that when you kind of take the eye off of public reporting that rate doesn't go down over time, is really where you want to be.

So, does that constant, continuous public reporting on this measure at that high level of performance warrant it to be continued? Or should it be in reserve status, meaning at any point in time, if people raise concerns that over time that measure is creeping down as we are not keeping as close an eye on it, it can potentially be moved back

into full status.

CO-CHAIR TIRSCHWELL: Mary?

MEMBER VAN DE KAMP: So, does the reserve status then -- would they still qualify for a measurement that CMS or insurance or someone who is looking at payment for best practice, would it still remain in that check "have to have it" box, if it reserve?

DR. BURSTIN: It would still be endorsed, but clearly demonstrated to be in reserve status. So, I think it would be sort of used with caution, I think might be a way for end-users to potentially interpret it.

CO-CHAIR TIRSCHWELL: And just for clarification, NQF endorsement doesn't mandatorily change policy of any external body, CMS, ASA, anything like that.

MEMBER VAN DE KAMP: Right. I understand that, but I just think if we are looking at down the line, which we all know they are looking at some sort of pay-for-

performance measures at some point, and this group is a strong group. I am worry that we take off something that is a strong process that is proven in health outcomes, that you don't get credit for. Back to your point, is it worth measuring because you are all doing it?

I don't know. To me, it seems like, as we look at this, you have to look at the broader base for ultimately what is a quality organization that is going to be looked at as a basis for making decisions. And if we take things off because they are already a gold standard, shouldn't you be rewarded? And isn't there some concern that, if you take it off of the gold standard and the checkbox, that it won't be, then, considered later as a best practice?

DR. BURSTIN: I think those are all valid concerns, but there are real opportunity costs to measurement. People tell us routinely, when they have a lot of things to

measure, you know, they are not measuring something else that might be new or particularly important.

Again, I think this is why it is a judgment call on the part of the Steering Committee.

CO-CHAIR TIRSCHWELL: Go ahead, Mike.

MEMBER KAPLITT: And performance gap is there for a reason, right? I mean, presumably, there are a lot of things we do every day that we don't measure anymore because everybody does it and, as you said, it is not necessarily worth continuing to do that.

So, the question is, what can NQF tell us in other areas has happened in the past to things that have been placed in reserve status? How have they been used? Does it really matter?

DR. WINKLER: I think it is a little too soon to tell. Reserve status was established just a little over a year ago for some of the

measures that CMS and The Joint Commission used around cardiovascular conditions. So, those measures have had a long history of being publicly-reported. They are at high levels.

One thing we have seen is that some of these measures CMS is not using in their value-based purchasing model, because when you are topped-out, the math doesn't work real well.

So, I think we are just beginning to see the implications of reserve status. So, we don't have enough of a history to really tell you how that is going to play out.

CO-CHAIR TIRSCHWELL: Go ahead.

DR. SCHWAMM: Yes, I just wanted to make two points. I mean, I think these are all really important and valid constructs. I just would make two observations.

One is that, with 800,000 strokes a year and rising, we are talking about a 4-percent gap is still over 20,000 patients a year. And

because all the exclusion criteria are there, we are still talking about 4 percent of patients who could be benefitting from this therapy for which we have some of the strongest evidence and the largest attributable benefit. So, I think that is incredibly important.

I think the other thing is that I do worry that, if it goes into this reserve status, sites will not be forced to collect on it. So, we will miss potentially an opportunity to understand if there is a larger performance gap out there than we have measured from this high-performing group of hospitals. If its NQF status is changed, it is possible that CMS and Meaningful Use will drop it as a measure, since they are not currently requiring its collection.

So, one alternative approach would be to re-endorse it for this cycle, give yourself a chance to see if, in fact, it really is topped-out nationwide. And then, it seems

like it would be very reasonable to consider it for reserve status, if, in fact, you see that that evidence is consistent.

CO-CHAIR TIRSCHWELL: Thank you.

Any other comments? Yes, go ahead.

MEMBER SULLIVAN: If I understood correctly, the question to the developer earlier about how the data was analyzed with regard to stroke center, hospitals, and those who are not, that that hadn't been looked at. So, it would seem like the next cycle might be an opportunity to look at that, and maybe there is, indeed, a bigger performance gap than we think, which an opportunity to look at that seems like something we ought not dismiss easily.

CO-CHAIR TIRSCHWELL: They do have the data now, but we don't have it on which to vote.

DR. BURSTIN: They haven't analyzed it, but you have data potentially to bring to this question.

MS. WATT: I don't know if in our Joint Commission ORYX system -- I know that we don't differentiate a primary stroke center hospital that is reporting from another hospital that is reporting. We could perhaps back into it because, of course, we know who all are certified.

But I think, though, the same argument applies. Hospitals are not required to collect core measure data. And so, the non-primary stroke center hospitals that have chosen to report as part of their accreditation requirement, they obviously have an interest in stroke care. And one could argue that they, too, are self-selected. They just haven't gone the step further to get the certification, and maybe that is why they are reporting it now, because that is their goal. I don't know that.

CO-CHAIR TIRSCHWELL: Okay. So, any other discussion? Fred?

MEMBER TOLIN: I was actually about to

take it down (referring to name tag).

But I think the comment that was made about looking perhaps reserve status next cycle I think seems very reasonable. No measure is going to reach 100 percent. There is always going to be a gap, for a variety of reasons that have already been elucidated.

And the other perspective I like to take on this is the burden that is placed on facilities in gathering the data. If a particular measure is doing well consistently, then I am very much in favor of the idea of moving that off the plate, so to speak, and looking at other measures that we can go forward with.

As Michael pointed out, we do things on a daily basis we don't measure because that is how we practice medicine.

CO-CHAIR TIRSCHWELL: Okay. Thank you.

So, as a process measure, do we need to take another vote? It is up to me? Or the group?

DR. BURSTIN: It is the will of the Committee. Would you like to entertain this measure for reserve status or do you want to just let your prior decision stand?

CO-CHAIR TIRSCHWELL: The prior decision being what?

DR. BURSTIN: You indicated it was suitable for endorsement. It was just your question as to whether -- do you want to take this next vote? It is just truly the will of the Committee. It, otherwise. does meet these criteria. It is, otherwise, evidence-based, highly-valid, reliable, et cetera.

Anybody can decide you want to vote. You don't have to vote to vote. But it is just, after this discussion, do you still want to proceed with voting and reverse that?

CO-CHAIR TIRSCHWELL: So, I would suggest we vote on whether we should proceed to put this in reserve status or not. Otherwise, it is endorsed.

Yes, Gail?

MEMBER COONEY: If we don't put it in reserve status, it continues as a measure or it disappears?

CO-CHAIR TIRSCHWELL: It is endorsed.

MEMBER COONEY: It is endorsed.

CO-CHAIR TIRSCHWELL: I don't know. That wouldn't have been the way that I would have done it, but can you type it up on the -- oh, you don't have a special loader for this?

MS. THEBERGE: No, it is just this. So, if you vote yes, you are putting it in reserve. If you vote no, you are keeping it as regular endorsement.

MS. JOHNSON: And both are endorsed. Just one is with reserve status; one is not with reserve status.

CO-CHAIR TIRSCHWELL: Okay.

(Vote taken.)

MS. THEBERGE: We are at 21 responses.

CO-CHAIR TIRSCHWELL: So, it is fully endorsed, not in reserve status.

DR. BURSTIN: And just keep in mind, the

Steering Committee votes are fairly early in the course of our process. So, this will all go out for public comment, and I am sure there will be some. So, we will bring that back to you after the fact and continue to iterate on this one some more.

MS. THEBERGE: For the record, that is 10 yes and 12 no.

CO-CHAIR TIRSCHWELL: I think we are going to take a break at this point for lunch.

Okay. We need a couple of minutes for public comment. First, local, the people that are here, or on the phone.

(No response.)

MS. THEBERGE: Amy, can you open the line for public comment?

OPERATOR: Fine.

If you would like to ask a question or have a comment, please press *1 on your telephone keypad. We will pause for just a moment.

(No response.)

Again, that was *1.

(No response.)

There are no questions or comments on the phone lines at this time.

CO-CHAIR TIRSCHWELL: Anyone in the room have a comment?

(No response.)

Okay, then, I guess we are adjourned for lunch, which is supposed to be a half an hour. I guess we should at least limit it to that, if not sooner. So, resume at 1:15.

(Whereupon, the foregoing matter went off the record at 12:46 p.m. and resumed at 1:18 p.m.)

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

1:18 p.m.

CO-CHAIR TIRSCHWELL: The next measure that we are going to do is 0435, a Joint Commission Stroke Measure 2, Discharged on Antithrombotic Therapy. And Gail is going to guide us through that.

MEMBER COONEY: Yes, I sort of thought that we should do this right after the other one because it has got all the same questions and issues.

This is a Joint Commission process measure looking at the proportion of ischemic stroke patients prescribed antithrombotic therapy at hospital discharge. The data collection and everything is very much the same. The data for impact and importance is very similar and was also rated highly.

CO-CHAIR TIRSCHWELL: So, let's go ahead and vote, if we can, on impact, 1a, for discharged on antithrombotic therapy.

Hang on one second until the timer comes

up there. Okay.

(Vote taken.)

MS. THEBERGE: It looks like 20.

CO-CHAIR TIRSCHWELL: Are we missing somebody? Salina? Oh, there she is.

Thank you. All right, no debate there. Sorry.

Evidence?

MEMBER COONEY: The evidence is also very similar to the last measure. Again, high in quantity, quality, and consistency.

CO-CHAIR TIRSCHWELL: Okay. So, let's go ahead and vote on the evidence for discharged.

(Vote taken.)

MS. THEBERGE: We are still short one vote.

Twenty-two high.

CO-CHAIR TIRSCHWELL: Great. And then, I guess this will be the rub, the question of performance gap.

MEMBER COONEY: As with the last

measure, the issue here is in the performance gap. All the different entities, The Joint Commission, the Get With The Guidelines, Stroke, all showed very high levels of compliance, both at the time they began the measurement in 2009 -- and this one is even higher; it was 98.1 when they began and 99.2 in the last quarter reported. And The Joint Commission notes a 2-percent performance gap.

CO-CHAIR TIRSCHWELL: Any comments? Yes, Salina?

MEMBER WADDY: I do think it is important to point out --

CO-CHAIR TIRSCHWELL: Could you push your button (referring to the microphone)?

MEMBER WADDY: I do think it is important to point out that, for the performance gaps, it is not necessarily just overall and general population, but also if there is a disparity that has been identified. So, that could also be an important reason to measure it.

DR. WINKLER: Is there any data on disparities from any of our data sources?

MEMBER WADDY: Yes, there was. There is and there is a disparity.

MEMBER COONEY: There was data from their original presentation. Oh, I'm sorry, not on disparities, but on the performance gap there was a difference. I don't believe there was on the disparity.

MEMBER GIDWANI: I believe there were disparities.

MEMBER KAPINOS: I think either they or another measurer has linked to those multiple publications about the fact that some groups, I think of females and African-Americans, there are some ethnicities that are definitely not receiving antithrombotics as much as the others.

MEMBER GIDWANI: On page 5, you will notice the factual data. They say that in a study using the uniform data system from nursing homes in five states that Blacks or

African-Americans were significantly less likely to receive antithrombotic therapy. They adjusted for a variety of covariates and found they were 80 percent as likely as Whites to receive that therapy.

MEMBER WADDY: And also, I believe that there is evidence as well in regard to that there is an underuse of antithrombotics in stroke in those patients, but I will have to go through really quickly the data that is presented here.

MEMBER KAPLITT: Right. It is true that I think there is reasonable data presented that there are disparities in use, but that is not what the measure is, right? We don't know what that use disparity is due to. Nursing home residents may have disparities because different groups go to different quality level of nursing homes, and some are much better at executing than others. It may have nothing to do with whether there is a disparity in discharge prescription, which is what is being

measured here.

So, I don't personally find the data on general disparities as relevant to what is being asked here.

MEMBER COONEY: And also, this measure does not in any way tell us whether we are making improvements in those areas of disparity because it doesn't measure those elements. It doesn't measure race. It doesn't measure sex. It doesn't measure location.

CO-CHAIR TIRSCHWELL: I have a question for the developer. This is a Joint Commission measure. Does the Joint Commission have any specific data on disparities? I think the answer is -- it is not in the report as far as I know.

MS. KOLBUSZ: We do address it in another section, that and the performance gap for the disparities section, but we do not collect data on disparities for any of the measures at this point.

CO-CHAIR TIRSCHWELL: So, there is evidence of more global disparities. But, as far as this measure being collected by The Joint Commission, there is no evidence presented of disparities? Does anybody disagree with that?

MEMBER KAPLITT: Well, no. It is just that, in anticipation of Lee making the same argument he made on the last one, that maybe we should not -- whatever it is called -- reserve this because it would be nice to at least know. If you are not going to collect the data, then that argument would be harder to buy, right?

DR. SCHWAMM: Yes, I mean, I think the only other point I would make is that there is data published two years ago in Get With The Guidelines looking at disparities in the prescription of antithrombotic therapy at 48 hours and at discharge that suggest that African-American patients, even after adjusting for site-level characteristics that

included the percentage of minority patients cared for at those sites, still had systematic reductions in the use of those therapies. So, we do have some evidence in hospitals, which includes Joint Commission hospitals, but that is not Joint Commission data that was submitted with this application. It is cited, but it is not --

MEMBER KAPLITT: But if there is no evidence in the disparity in the actual prescribing of it, couldn't that argue --

DR. SCHWAMM: No, there is evidence.

MEMBER KAPLITT: But I thought you said the evidence was that, despite all that, there is evidence of a disparity in use afterwards?

DR. SCHWAMM: No, no, this is all about hospital-based prescribing. It is all relevant to these measures. My point --

MEMBER KAPLITT: So, there is actually a gap in hospital-based prescribing upon discharge, and there is data for that. We just don't have it?

DR. SCHWAMM: Right. So, the gap is that there is a differential use among minority patients. So, there is a differential performance there -- we do know that -- even though overall performance is high.

CO-CHAIR TIRSCHWELL: Right, and that is sort of indirectly supportive of the disparity, not exactly, in The Joint Commission measure, but largely overlapping, we would assume?

DR. SCHWAMM: Correct.

MEMBER KAPLITT: Well, okay. I mean, I am not understanding because this isn't use. This is prescribed, right? You are saying there is data -- I just want to be clear.

Because I could make the argument that, if the data showed that the prescribing was equivalent but the use wasn't, that this is an irrelevant measure.

DR. SCHWAMM: We don't have any data in Get With The Guidelines on use after

discharge.

MEMBER KAPLITT: Okay. So, this is specifically talking about prescribing, not use, that there is disparity, that there is data on disparities?

DR. SCHWAMM: Exactly.

MEMBER KAPLITT: Okay.

DR. SCHWAMM: I mean, there was also data on use, but that is not the scope of this.

MEMBER KAPLITT: Right. Okay.

CO-CHAIR TIRSCHWELL: Jack?

MEMBER SCARIANO: I think that what Mike was just asking the actual same question. If you just write for them or if you just tell them how many people get it, there is a cost factor. Also, patients worry about bruising.

I mean, in my clinical practice, I see in the Medicaid population probably only about half of them get it. Aspirin at, say, Walmart costs, it is about $10 a bottle for just one month's worth of 81 milligrams. So, a whole

lot of them can't afford it.

CO-CHAIR TIRSCHWELL: I mean, that is an important issue and use after discharge is, of course, where the rubber hits the road, but it is not what is being measured here. So, I don't think we can tackle that issue today.

Any other comments about the performance gap? Gail?

MEMBER COONEY: I don't know if this is where it belongs, but they do make the comment that there is no evidence that either higher aspirin dose or any other antiplatelet regime was more effective than low-dose aspirin. If we are trying to increase compliance, I mean, could we just measure why people didn't get aspirin, aspirin unless otherwise specified, and make this more useful?

CO-CHAIR TIRSCHWELL: I guess I am not sure how that addresses the performance gap question.

MEMBER COONEY: It doesn't. Never mind. Yes.

MEMBER SCARIANO: Well, actually, 81-milligram aspirin is a whole lot higher than just regular aspirin.

CO-CHAIR TIRSCHWELL: Price-wise?

MEMBER SCARIANO: Right.

CO-CHAIR TIRSCHWELL: Yes. Okay. Extraneous issues.

So, you know, I think we have to vote on whether we think there is a performance gap specifically about the prescription at discharge, not the use after.

And let's go ahead and open the voting.

DR. BURSTIN: And just to be clear, performance gap includes variation between providers or populations, just to be clear.

MEMBER KAPINOS: But it actually goes in the opposite direction. Maybe in the future you should break it into two subgroups, right? Because if you have a low-performance gap with a lot of them compliant, but disparity, then you don't know how to vote on that item.

(Vote taken.)

MS. THEBERGE: We are at 20 votes.

Five high, 11 moderate, 5 low, 1 insufficient.

CO-CHAIR TIRSCHWELL: Okay. So, we will proceed along. We are on 2a, reliability.

MEMBER COONEY: Thank you. I wasn't turned on (referring to the microphone).

Again, there is the comment about confusing presentation of the reliability data. That is the same issue with the formatting. But, in general, the data was in agreement as The Joint Commission has been presenting their data. And that is with the individual elements.

CO-CHAIR TIRSCHWELL: Any other comments or questions about reliability?

(No response.)

So, we will go ahead and open up the voting.

(Vote taken.)

MS. THEBERGE: Eighteen high, 4 moderate.

CO-CHAIR TIRSCHWELL: 2b, validity.

MEMBER COONEY: Again, the validity was originally assessed by a survey, and The Joint Commission also has ongoing validity testing based on consumer-use feedback.

CO-CHAIR TIRSCHWELL: Any other comments or questions about validity? Risha?

MEMBER GIDWANI: In this measure, and in other measure specifications, it notes that the face validity has been tested in previous surveys. In the future, it would be very useful if we could actually see the results of the original surveys as opposed to having that sentence only.

CO-CHAIR TIRSCHWELL: Noted.

Any other comments or questions?

(No response.)

Let's go ahead and vote on validity.

(Vote taken.)

MS. THEBERGE: Fifteen high, 7 moderate.

CO-CHAIR TIRSCHWELL: Okay. Moving on to -- what is it, feasibility? -- no,

usability.

MEMBER COONEY: Under usability, the data is available on The Joint Commission website, and is also used for maintaining their primary stroke center designation.

And then, the question came up as to how usable is it if there is no performance gap.

CO-CHAIR TIRSCHWELL: Right.

Questions or comments? Risha?

MEMBER GIDWANI: I will ask the same question I did before. And that is whether the CMS Meaningful Use specifications are the exact same as The Joint Commission's specifications.

MS. WATT: The answer is the same answer. They are intended to be, but they have not been tested.

CO-CHAIR TIRSCHWELL: Any other comments or questions?

(No response.)

Let's go ahead and vote, then, on usability.

(Vote taken.)

MS. THEBERGE: Fourteen high, 8 moderate.

CO-CHAIR TIRSCHWELL: And then, finally, feasibility.

MEMBER COONEY: Many of these elements, all these elements are generated during care process. Many of them are captured in electronic data, but, generally, they require paper as well. But they are already in use successfully.

CO-CHAIR TIRSCHWELL: Any other comments or questions about feasibility?

(No response.)

Let's go ahead and vote then.

(Vote taken.)

MS. THEBERGE: Fifteen high, 7 moderate.

CO-CHAIR TIRSCHWELL: And then, finally, overall suitability for endorsement. Any questions or comments before we vote on that?

MEMBER COONEY: You know, I think this comes up to the same questions we dealt with

on the last one of an ongoing measure with higher performance.

CO-CHAIR TIRSCHWELL: So, the only thing you are suggesting might mitigate your endorsement might be the high performance --

MEMBER COONEY: Correct.

CO-CHAIR TIRSCHWELL: -- that already exists?

Anything else?

(No response.)

So, let's go ahead and vote, then, for overall suitability.

(Vote taken.)

MS. THEBERGE: We are at 20. We are still short two votes.

CO-CHAIR TIRSCHWELL: Will everybody vote again?

MS. THEBERGE: Can everyone vote one more time? You do have to point it at me for the receiver to pick up. All right. There we go.

Twenty-one yes, 1 no.

CO-CHAIR TIRSCHWELL: Okay.

DR. BURSTIN: Just one point?

CO-CHAIR TIRSCHWELL: Yes, please.

DR. BURSTIN: So, part of the work we also do at NQF around disparities is trying to identify which measures are disparity-sensitive, primarily based on the impact, the gap in care, compared to non-disparity populations. And so, this just seems like, particularly from the perspective of The Joint Commission, a measure that would likely be designated as disparity-sensitive.

So, there are ways, I would hope, both with Get With The Guidelines as well as Joint Commission, to begin collecting the data, to be able to stratify it. It would certainly be required by the time of maintenance.

MS. WATT: Good point, and thank you.

Up until now, and even as we speak, we have not had the opportunity to look at disparities particularly based on race and ethnicity because these are not general data

elements that have been collected.

But we are adding them to our specifications now going forward. And so, because there is always a lag time, I think it is probably going to be a year before we start getting data on them. But, yes, we are working on it. Thank you.

DR. SCHWAMM: And just to clarify, the Get With The Guidelines has collected race and ethnicity data using the Census 2000 recommendations. We have just moved to adopt as an optional data element the more diverse race classifications underneath Asian and some of the other minority designations. So, we are hopeful that we would be able to provide a very high level of specificity around undertreated populations.

CO-CHAIR TIRSCHWELL: Salina?

MEMBER WADDY: So, I was wondering how does NQF plan to address that? I mean, none of your elements are really specific to disparities. So, for example, within this

question and the last one, whether or not it would go in reserve, it might go on reserve for the general population, but it may be of value for specific populations. And so, is there an opportunity to make that recommendation to you?

DR. BURSTIN: So, just to be clear, we would not put something on reserve status where there are known disparities, because gap refers to either a total gap or a gap for different populations. So, we would actually see that as something that would not move something onto reserve.

That is why I added that point about its being --

MEMBER WADDY: Right. No, that is actually why I initially brought up --

DR. BURSTIN: Right.

MEMBER WADDY: -- the point, when I was thinking about --

DR. BURSTIN: Absolutely.

MEMBER WADDY: -- to why we may consider

putting something on reserve status --

DR. BURSTIN: Right.

MEMBER WADDY: -- but also may consider not putting it on.

DR. BURSTIN: So, we currently have a Disparities Committee that has been going through and trying to really codify exactly how we will determine measures that are disparity-sensitive.

And what we will do prospectively, probably beginning in about six months, is have every committee go through an exercise after you say suitable for endorsement to indicate whether or not it should be disparity-sensitive and always recommended to be stratified.

So, we have done a retrospective analysis of about 400 measures that the Committee is working through, but really doing this prospectively with the data on maintenance from the actual measures is really how you want to do it.

MEMBER WADDY: And I did want to mention just one more thing really quickly. I did go look back through the articles that are listed under the disparities section. They do list some of the articles. It is not purely whether or not a patient decides to take a medication or not, but there are differences as to whether or not physicians actually prescribe the medications. Those are within some of the articles, particularly in the REGARDS articles that are listed.

CO-CHAIR TIRSCHWELL: Thank you for those comments.

CO-CHAIR KNOWLTON: Okay. We are moving on to 0325. And David is presenting.

CO-CHAIR TIRSCHWELL: Yes. So, this is a similar measure. It is also called "Discharged on Antithrombotic Therapy," but there are a couple of differences compared to the previous measure.

One of the differences is that in the denominator of this measure it includes not

only patients with a diagnosis of ischemic stroke, but also with a discharge diagnosis of TIA.

And another big difference is that, as opposed to The Joint Commission measure, which was a facility-based measure, this one has the level of analysis being the clinician. But, otherwise, in intent and purpose, it is highly similar.

So, then, just going to impact, I think the Work Group was unanimous in endorsing a high level of impact.

Anybody want to make any comments? Or shall we just proceed to voting?

Sorry, Salina, go ahead.

MEMBER WADDY: I think this also includes rehabilitation, right, performance indicators within the rehabilitation center? And there can be a lot of differences between patients, what their discharge medications versus they receive in rehab versus what they receive in the outpatient setting that can be

really important.

But, I have been trying to go through that really quickly right before vote. So, I just think it is an important indicator if it also captures that element as well.

CO-CHAIR TIRSCHWELL: So, you are pointing out that rehab inpatients would also be captured under this measure where --

MEMBER WADDY: Yes, at least within the measure title it includes the stroke and stroke rehabilitation. So, I am just trying to understand what the setting exactly is. Is this also within a rehab setting; whereas, the other is discharged from the hospital?

CO-CHAIR TIRSCHWELL: I think it is possible that all of the AMA measures are called "stroke and stroke rehabilitation".

CO-CHAIR KNOWLTON: Yes, let's ask the developer what they --

MEMBER WADDY: So, this doesn't -- sorry about that. Go ahead.

MS. JOSEPH: So, the measure is actually

only in the inpatient hospital setting.

MEMBER WADDY: Okay.

MS. JOSEPH: The title "stroke and stroke rehabilitation," just for your knowledge, is the title of our actual whole enter performance measurement set.

MEMBER WADDY: Okay.

CO-CHAIR TIRSCHWELL: Okay.

MEMBER WADDY: Then, never mind.

CO-CHAIR KNOWLTON: Then, are we ready for impact?

(No response.)

Vote, please.

(Vote taken.)

Okay.

MS. THEBERGE: Twenty-two high.

CO-CHAIR KNOWLTON: Okay. Evidence, Dave?

CO-CHAIR TIRSCHWELL: So, evidence, again, was really the same evidence that has been presented for all of the other measures related to the effectiveness of antithrombotic

therapy in the prevention of subsequent vascular events. And so, it met with really pretty universal acceptance.

I should say, yes, I don't have anything to add.

CO-CHAIR KNOWLTON: Anybody have anything to add?

DR. BURSTIN: This measure also includes TIA, in addition to stroke. So, it is not the same patient population; it is broader.

DR. WINKLER: Right, and so, the question is, is the evidence equivalent for both of those groups?

CO-CHAIR KNOWLTON: Anything further?

CO-CHAIR TIRSCHWELL: Nothing further from me.

CO-CHAIR KNOWLTON: Any questions from the group?

(No response.)

Okay. Let's vote.

(Vote taken.)

MS. THEBERGE: Nineteen yes, 3 no.

CO-CHAIR KNOWLTON: Performance gap.

CO-CHAIR TIRSCHWELL: So, performance gap, this was quite different than the previous measure. Despite the fact that The Joint Commission showed extremely high levels of compliance, this performance gap as presented suggested a number of very high rates of non-compliance.

At one point, they quote 53 percent of the patients did not meet the measure. In other places, it says that more recent data suggested 80-plus percent compliance. So, there is certainly a mismatch between what these guys are measuring, whether it is at the clinician level. I am not sure what the difference is, what explains that big gap in why they are seeming like there is a big gap here and not at the facility level.

I guess, if it is okay with the Chair, I would ask the developer if they have any comments on that.

CO-CHAIR KNOWLTON: Sure.

MS. JOSEPH: Sure. So, the performance gap data that we have included is actually from the PQRS program. And again, as you stated earlier, this is at the physician level of measurement.

So, the data that we included shows that in 2008 the performance gap was 53.03 percent; in 2009, 85.1 percent, which shows a great improvement, and then, 2010 was 82.8 percent. And that is the most recent data we have from the PQRS program.

CO-CHAIR TIRSCHWELL: And again, I don't understand -- it sounds like a huge gap to me, but --

MEMBER WADDY: Yes. Do you have any evidence as to whether or not it is because you included the TIAs, and some of these TIAs were more spells rather than clear-cut TIAs without ABCD-squared scored or anything?

MS. JOSEPH: We do not have that information.

MEMBER WADDY: And maybe they didn't

treat those?

MS. JOSEPH: We don't have the information with regard to the cause and why the performance rate is what it is in the PQRS program.

CO-CHAIR TIRSCHWELL: So, I mean, in contrast to The Joint Commission measure, the data, as presented, suggests a large performance gap.

DR. BURSTIN: It is probably worth looking at the 2009-2010 data, which is more like 85 percent, rather than the --

CO-CHAIR TIRSCHWELL: Still, relatively speaking, that is pretty large.

CO-CHAIR KNOWLTON: Go ahead, Amy.

MEMBER BARRETT: I would make a comment that this is physician-level data. So, these would be people practicing, presumably, physicians practicing in all kinds of settings, right, stroke centers and other centers as well?

CO-CHAIR TIRSCHWELL: Thank you.

CO-CHAIR KNOWLTON: Mary?

CO-CHAIR TIRSCHWELL: The same comment? Michael?

MEMBER KAPLITT: Can you just clarify at least for us -- I mean, I understand where the data came from -- but how you generated this? What exactly are you defining as a performance gap within this data? This is confidential data. So, the question is, what is your measure?

MS. JOSEPH: Sorry. The 2008 data is confidential. The 2009-2010 data is not confidential. It is actually posted on the CMS/PQRS website.

The way that the 2008 data is presented, it is by percentile, exactly as we dropped it into the form. So, the way they report it, there is a gap in care shown by this data as 53.03 percent of patients reported on did not meet the measure. All we have is the performance rate to report.

CO-CHAIR KNOWLTON: You have got to use

your microphone, please.

MEMBER SCARIANO: Yes, is this based on actually what you found in the hospital charts or is it like an actual questionnaire that someone fills out?

MS. JOSEPH: The PQRS system is currently a volunteer reporting program. And so, for the 2009 and 2010 Patient Experience Report that is cited here to support that data, it said that about 24 percent of eligible professionals participated in 2010. And so, the performance rates may not be nationally-representative.

MEMBER SCARIANO: Well, that is a question on this point. So, I mean, we haven't gotten to that point, but, presumably, their exclusions or exceptions, or whatever you guys define, are those the same as what PQRS defines? Do you know? Is PQRS simply measuring did you do it or didn't you; whereas, you guys have all kinds of exceptions, in which case, then, the

performance gap would be very different?

MS. JOSEPH: So, the PQRS program, also, we do have the exception rates that are reported as well as the performance rates in the PQRS program, and they are the same.

CO-CHAIR KNOWLTON: Any other thoughts, questions, comments on this?

(No response.)

Anything to add, David?

CO-CHAIR TIRSCHWELL: No.

CO-CHAIR KNOWLTON: Okay. Let's vote on the performance gap then.

(Vote taken.)

MS. THEBERGE: We have 13 high, 6 moderate, and 3 insufficient.

CO-CHAIR KNOWLTON: Okay. So, let's go on to reliability.

CO-CHAIR TIRSCHWELL: So, for reliability, it seems the same dataset as was used for the facility-level reliability earlier was used for this clinical-level reliability analysis here. And they did chart

re-abstraction, and the kappa statistics were very high. So, it seems like there is pretty good evidence that this information is being reliably-obtained as far as the evidence presented.

CO-CHAIR KNOWLTON: Questions or comments? Greg?

MEMBER KAPINOS: But that is where we had the issues about the TIA definition, right? That would be impacting the reliability, correct? Or should we keep that coming for a little bit later?

CO-CHAIR TIRSCHWELL: I think the reliability refers to whether what they are getting first-pass from the charts is then recreated by an independent abstractor getting a second-pass. So, they are sort of identifying the patients already. It doesn't really address the reliability of the diagnosis of the TIA or ischemic stroke.

CO-CHAIR KNOWLTON: Therese?

MEMBER RICHMOND: Can I just go back?

So, it is the same dataset that is being used, but here it is called the clinician level, and previously it was called the facility level. So, I am confused in terms of what I am looking at. Could somebody clarify that for me?

CO-CHAIR KNOWLTON: Developer? Did you have a question?

MS. HANLEY: Hello?

So, yes, this is the same dataset. This was tested at the physician level. There are four sites. Again, we don't report out performance rates by site. However, this is reported at the physician level. It is an inpatient measure, though.

CO-CHAIR KNOWLTON: Go on with your question, Therese.

MEMBER RICHMOND: I am not sure how else to ask it. When I look at the data sample, it looks identical to me. So, I just want to verify that, from the same sample, sometimes it is reported at the provider level and

sometimes it is reported at the facility level? It just does not seem logical, I guess is where I am.

CO-CHAIR TIRSCHWELL: I think I might be able to partially address that. I think it says in a couple of different places that this reliability testing was done at the element level. So, they were looking at the reliability of the data elements. And then, sort of the logic of how it falls out, whether it is facility level or clinician level, is sort of subsequent to the reliability of the assessment of the individual data elements.

And so, I think that is how they were able to use the same dataset, because they were really just looking at individual item reliability. And you know, there are 20 items that go into the logic for each one of these things. It was just the reliability of the individual items that they were assessing. Is that correct?

MS. HANLEY: It really comes down to how

the measure results are aggregate.

CO-CHAIR KNOWLTON: Any other questions on reliability?

(No response.)

Okay. Let's vote.

(Vote taken.)

MS. THEBERGE: We need one more vote.

Tied at 11 high and 11 moderate.

CO-CHAIR KNOWLTON: Validity?

CO-CHAIR TIRSCHWELL: So, validity -- and this is, in the form it has the same face-validity evaluation that the others do where the experts voted on whether they thought it was a valid measure. You know, they strongly agreed in most cases and a few just agreed without strength. But that was the data presented.

CO-CHAIR KNOWLTON: Jolynn, you had a comment on this issue?

MEMBER SUKO: Yes, I am really struck by the difference between this and The Joint Commission measure, considering that even with

the TIAs, we are measuring, essentially, similar concepts.

And I know the PQRS program has two different methods for measure submission, one which is a database method, which is, I think, based largely on chart abstraction, and one which is a super-bill method. At least that is the way it was two years ago; it may have changed.

And I am wondering, has there been any testing done to compare these data? I mean, is it a function of how we are submitting these data that we are getting such what appear to be differences in our impact on this?

I mean based upon this face validity, I am having a really hard time understanding the differences.

CO-CHAIR KNOWLTON: Do you want to comment or will I go to the developer? Developer, do you want to comment on that?

MS. HANLEY: Sure. So, CMS it the

primary implementer of this measure. And you are correct, you can either report through a claims-reporting mechanism, where the individual physician reports whether or not it was prescribed or whether or not there is a valid exception for that patient, or they can work with a registry that actually works with the practice, for the individual physician to aggregate the results and report the aggregate results on behalf of CMS.

MEMBER SUKO: So, to me, that begs a question of, in your testing, are you testing that claims-based method against the actual chart?

MS. HANLEY: The answer to that question is, yes, when we could. There are four of the measures that we were able to do parallel forms of reliability and do that comparison.

For this one -- give me one moment to find my place here -- this site was actually not submitting on this measure. So, we weren't able to perform that test.

CO-CHAIR TIRSCHWELL: So, I would just add that somebody on the Work Group call brought up whether the low performance could just be related to non-documentation in claims data or something along those lines, artifactual of the data collection, as opposed to a true performance gap. And it sounds like it is not specifically answered for this performance measure.

And I mean, we are talking about validity now, and 2b.1 there says, "specifications consistent with evidence". So, I think it is appropriate to bring up the TIA question again and see. It seems like there were some feelings about whether that should be included or not, and at least there was some comment about trying to harmonize it a little bit, which I think I am not supposed to say until later in the day, but with the other measures that don't include TIA.

Does anybody want to comment about the applicability of TIA?

CO-CHAIR KNOWLTON: Anybody have a comment on that? Oh, I didn't see it. I'm sorry. Salina?

MEMBER WADDY: So, I think what would be potentially helpful is if you had two separate measures, one that was simply stroke and then one that was simply TIA, instead of just this stroke and then stroke plus TIA, if you really wanted to know how people with TIAs are actually being treated.

CO-CHAIR KNOWLTON: Anybody else? Dan?

MEMBER LABOVITZ: I am one of the Work Group members with feelings.

(Laughter.)

I have actually done some work on diagnosis of TIA. The validity of the diagnosis is widely known to be problematic. What I have been seeing is wide ranges of ratio of TIA to TIA-plus-infarct across institutions, between states, between hospitals that have neurology house staff and those that don't.

You also have a problem with billers who will look for any mention of TIA and then that becomes the diagnosis. If somebody writes, "Rule out TIA," that generates a lot more money than carpal tunnel. So, it is a problem, but it is a problem that I think needs to be fixed.

But I think if we don't measure TIA, if we don't include it, and if we don't hold hospitals' feet to the fire, we will never fix it. I think TIA, if we abandon it, will remain a garbage pail forever.

CO-CHAIR TIRSCHWELL: And I would just reiterate that point. I think that is a great point.

If, in fact, the billers are somewhat inappropriately moving the TIA to the primary diagnosis, when in fact the clinicians really didn't think so -- and God knows that clinicians have nothing to do with the discharge coding, as far as I can tell -- then this would theoretically identify those

hospitals, and they would suddenly be performing lower on this measure. And then, it would all come around the horn and, hopefully, improve that issue.

CO-CHAIR KNOWLTON: Salina?

MEMBER WADDY: I definitely agree with that point, but I do think it is important to be able to tease apart those two things. That is why I think having a separate measure for TIA that is just as rigorous and emphasizes it just as much would be important.

MEMBER COONEY: This is not, as I understand it, a facility-based measure. So, I don't think it is going to tell us anything about how a hospital's data abstractors recognize or don't recognize TIA, because this is coming off physician claim forms.

CO-CHAIR TIRSCHWELL: I guess it depends upon how your physicians generate their bills and all sorts of variations, although, yes, well, that is a good point.

MEMBER KAPINOS: Actually, about that,

like teach me something like -- I was told, and I am a junior attending, and for my first year of billing, actually, subarachnoid hemorrhage, when you have vasospasm, I bill actually exactly the code of a TIA, I was told, for vasospasm-induced. So, it is transient or sometimes permanent, but not necessarily resulting in an infarct. It is ischemia due to vasospasm after a subarach.

So, I may be participating greatly in that big mess.

(Laughter.)

So, is there another code?

CO-CHAIR TIRSCHWELL: So, for your inpatient strokes, I am sure that your primary diagnosis is a subarachnoid hemorrhage which has got to reimburse at a DRG rate much higher than a TIA.

You know, that is all speculation, and it is probably not the main point here.

CO-CHAIR KNOWLTON: Do you have a point?

DR. BURSTIN: Yes, I would just make a

point. So, it does appear, though, that the denominator is still based on a hospital discharge ICD-9 code. So, that is still the denominator; is that correct? It appears to be.

And if so, one potential option would be that you could recommend that the measure have two rates as part of it, one for stroke, one for TIA.

DR. DROZDA: This is Joe Drozda. Could I say something? I am one of the measure developers?

CO-CHAIR TIRSCHWELL: Yes, please go ahead.

DR. DROZDA: Yes, I am Co-Chair of the PCPI Work Group. We developed this measure.

As far as the two measures go, I am finding myself taking the opposite side of an argument that we had during discussion of the chronic coronary artery disease measures where there were two measures up for consideration. I believe this was an aspirin measure, and I

think our chronic coronary disease measure set called for the denominator to be patients with chronic coronary disease. And there was a competing measure that had the denominator as patients with ischemic vascular disease of all types.

And ultimately, the broader measure was chosen because of NQF's desire or principle to go for broader populations rather than narrower populations. We find ourselves going the other direction here where we have actually put in this measure a broader population.

So, I think what I am asking for is some sort of consideration of consistency across measure sets, because I rather eloquently made the argument that was just made for dicing out patients with coronary disease previously and lost that argument. So, I am a little bit personally conflicted, but I think we are trying to be consistent here and choose the broader population.

DR. BURSTIN: And, Joe, this is Helen.

Just to respond, I completely understand your point. And our preference, obviously, is that measures should be as broad as the evidence is appropriate. And so, in this case, it has already passed through the evidence category with a clear sweep that it is fine to include both stroke and TIA. The only questions, I think, were up to validity. If there are concerns about the validity of the coding of TIA, then you want to be able to, in fact, see those differences.

I think the broad population is great. The question is, is it possible to just report two rates?

DR. DROZDA: And if we find ourselves arguing over validity of coding, I mean I don't if we will ever be able to get any measures through because we run into that everywhere. We have to rely on the validity of the coding. How else can we do this?

CO-CHAIR KNOWLTON: Dan?

MEMBER LABOVITZ: I am feeling everybody's pain.

(Laughter.)

I think validity of coding is a problem, congestive heart failure perhaps the worst area. We suffer in a lot of places. But it is possible to make it better.

I work with our billers. We go back and forth. Every diagnosis gets reviewed.

And I think hospitals will be pressured because we are using hospital data, even if it is individual private providers doing it, we are using hospital ICD-9 codes, I think hospitals are going to be pushed to perform better in deciding, yes, that really wasn't a TIA. They have a financial incentive right now to diagnose it every time possible. Now they are going to have another incentive to perhaps dial back a bit.

DR. DROZDA: And I really like that argument. I resonate with that strongly. And I think that is even true at the physician

coding level, if there are concerns about that. I think this puts discipline in coding.

CO-CHAIR TIRSCHWELL: Bill, go ahead.

MEMBER BARSAN: Yes, that idea of splitting these out, making them be reported separately, is great. When does that happen? I mean, do you go all the way through the process and then it comes out later as a recommendation? Or where does that, how do you do that?

DR. BURSTIN: Some of that could just be a recommendation that you would potentially make to the developer. Again, we are probably getting a little bit ahead of ourselves. You guys raised the issues of harmonization to The Joint Commission measure. This might be something best saved for later when we talk about comparability to the other measure.

CO-CHAIR KNOWLTON: Yes, are we ready to vote on validity?

(No response.)

Okay.

(Vote taken.)

MS. THEBERGE: We have 9 high, 7 moderate, and 6 low.

CO-CHAIR KNOWLTON: Okay. So, now we move on to usability.

CO-CHAIR TIRSCHWELL: Usability in the Work Group, there was a little bit of mixed feeling about it, but no strong objections at all. And I don't have any particular reason to argue against it.

CO-CHAIR KNOWLTON: Anybody else have a comment in this arena?

(No response.)

Okay. Let's vote on usability.

Oh, yes, Salina? Sorry.

MEMBER WADDY: I think that the usability question gets back to the same topic that we were discussing regarding the inclusion of TIA, how do you define the TIA, how important or how really clean that definition is within the hospital system setting and ICD-9 coding.

CO-CHAIR KNOWLTON: Dan, is that a residual card up on your desk or --

MEMBER LABOVITZ: Yes.

CO-CHAIR KNOWLTON: Okay.

MEMBER LABOVITZ: I was enjoying all the attention.

(Laughter.)

CO-CHAIR KNOWLTON: Anything else?

(No response.)

Let's vote on usability.

(Vote taken.)

MS. THEBERGE: Twelve high, 8 moderate, 2 low.

CO-CHAIR KNOWLTON: Okay. Feasibility.

CO-CHAIR TIRSCHWELL: Again, not much to report here. I mean, it has been used. It has been collected. It seems feasible.

We had some discussion in our Work Group -- I don't know if it was this one or one of the other PCPI ones -- about claims, that it is all in the electronic health record, and there was some ambiguity around that. But I

think at the end it could be implemented electronically.

CO-CHAIR KNOWLTON: Yes, Risha?

MEMBER GIDWANI: I see in the measure specifications that there is both ICD-9 or ICD-10 codes as well as CPT codes for the denominator. I am just not really familiar with, do physicians have access to both or is it generally that they have the CPT codes and the hospital has the ICD-9 codes?

CO-CHAIR KNOWLTON: Helen, can you answer her question, so she can hear you?

DR. BURSTIN: Well, I mean, the clinician has access to both, but the hospital would have ICD-9. CPT is uniquely, my understanding, outpatient.

CO-CHAIR KNOWLTON: Any other questions on feasibility?

(No response.)

Okay. Let's vote.

(Vote taken.)

MS. THEBERGE: Fifteen high, 7 moderate.

CO-CHAIR KNOWLTON: Okay. And we move on to overall suitability for endorsement. David?

CO-CHAIR TIRSCHWELL: I don't have anything further to add. It was unanimously approved in the Work Group.

CO-CHAIR KNOWLTON: Other thoughts?

(No response.)

Okay. Let's vote.

(Vote taken.)

MS. THEBERGE: Twenty yes, 2 no.

CO-CHAIR KNOWLTON: Okay.

CO-CHAIR TIRSCHWELL: All right. We have to jump out of order a little bit because one of the developers needs to leave. So, I think it is 2017 that we have to do next. So, we are going to temporarily skip over 0439.

And so, 2017, Ramon, can you --

MEMBER R. BAUTISTA: Yes. Thank you, Mr. Chairman.

Yes, we are going to be discussing the use of CT scans and MRI reports. So, the

study numerator is the final report of initial CT scan or MRI that includes documentation of the presence or absence of each of the following, hemorrhage, mass lesion, and/or acute infarction, while the denominator are all the final reports of CT scans and MRIs of the brain performed either in the hospital 24 hours of arrival or in the outpatient imaging center to confirm diagnosis of stroke, TIA, or intracranial hemorrhage for patients more than 18 years or older with a diagnosis of ischemic stroke or TIA or intracerebral hemorrhage.

There was a healthy discussion during the Work Group. So, I may need the help down the line of my colleagues from radiology, stroke, and maybe even the non-neurologists in the group.

There were three Committee members who actually evaluated this during our Work Group meeting. We were looking at the impact of 1a. There was one high, one medium, and one low score.

And I think part of the reason was there was maybe at that point a little bit of an uncertainty if we were actually looking at importance of the measure or scientific acceptability.

Some of the comments given, though, were that while stroke affects large amounts of people, the link between neuroimaging and being able to improve outcomes was not well-established. Some people in the group felt that we are not sure if the issue was the lack of documentation or the lack of actual imaging that took place. That wasn't really clear in the initial measure. So, that was how that turned out.

CO-CHAIR TIRSCHWELL: Anybody want to comment on the impact? Salina?

MEMBER WADDY: I just have a question to the people who developed the measure, why mass lesion was included.

CO-CHAIR TIRSCHWELL: Is the developer here?

DR. SEIDENWURM: Yes, sure.

CO-CHAIR TIRSCHWELL: Go ahead.

DR. SEIDENWURM: The reason we included that is for two reasons. Mass lesion was one of the exclusions in the NINDS trial and, also, the secondary analysis of the ECASS and the follow-on ECASS showed that the presence of mass effect was a predictor of adverse outcome from t-PA administration. So, that is why we included that.

MEMBER WADDY: So, is it mass effect and not necessarily mass lesion? Is that what you were asked?

DR. SEIDENWURM: Well, both.

MEMBER WADDY: So, it could be a brain tumor?

DR. SEIDENWURM: Yes, that is a contraindication to the administration of t-PA.

CO-CHAIR TIRSCHWELL: It seems to me a lot of what drove some of the contents -- and correct me if I am wrong -- is sort of

contraindications to IV t-PA and imaging reports. And so, I think those were some of the other things that could imitate an acute stroke, although I would hope a slow-growing brain tumor wouldn't mimic an acute stroke, but there are times when they bleed, or whatever.

Bill?

MEMBER BARSAN: Yes, I am just curious, I mean, if this is supposed to get back to something to do with t-PA treatment, my experience is usually the radiology readings come up hours or days after you have treated somebody with thrombolytic therapy, and it is really of no value to the person actually doing the treatment. I mean, you rarely, if ever, get a read by that. I mean, what is the point? I'm not sure.

Does the developer want to respond to that? Or, actually, why don't we let Dr. Hackney respond first?

MEMBER HACKNEY: Well, I agree that I am

not enthusiastic about this as a measure, but if you are a stroke center, you are not supposed to be generating your reports hours and days after the study is done.

But the concern I have is what difference it makes. It matters whether you have a mass lesion, for example, and that is correctly diagnosed. But I don't think it matters whether you say that there isn't one when it is true that there isn't one if you said the study is normal. Without saying anything specifically about mass lesion, you would have conveyed the same information. So, I am not sure that it makes sense as a performance measure. We have our residents do it to remind them of what they are looking for, but not because we think it has any influence on patient care.

CO-CHAIR TIRSCHWELL: And I think that is a great point. I would have one other comment about the specifications, but it is sort of an overall question.

They also say that they are including patients with one documented symptom consistent with ischemic stroke or TIA or intracerebral hemorrhage. And I guess I have no idea how you even identify who the people are that had one of the symptoms for that. Can the developer comment on that?

DR. SEIDENWURM: Well, there are two questions that have been asked. I will address the first one, and then staff will take the second.

With respect to the need for the correct documentation, I think the point that we are getting at here is a clear and unambiguous interpretation of a CT or MRI scan that can be relied upon rapidly in the emergency room. And it is particularly in settings where there aren't residents, where there aren't stroke teams, where there might be an emergency room doctor with telephone neurology backup or no neurology backup.

So, the point is to remove the

uncertainties from the reporting of the CT scans and to encourage timely reporting of these necessary data elements. So, I think that is what we are getting at here. That is the nexus between this type of a performance measure and the improvement in stroke care.

MS. HANLEY: And to the point about the symptoms, as Dr. Seidenwurm mentioned, the focus is to address these areas on the report, which is just as important to do if the patient actually has a stroke or a hemorrhage as if they don't. And so, it includes those CTs or MRIs for patients who come in with symptoms that might be suggestive of stroke, but aren't actually, then, diagnosed within the time --

CO-CHAIR TIRSCHWELL: And I apologize. I don't have the details. So, you are saying that the denominator is defined by the symptoms that are written on the top of the radiology report?

MS. HANLEY: The denominator is defined

with the procedure, CT or MRI, and then those studies that either have a final diagnosis of stroke or hemorrhage or TIA, but, also, for those, essentially, they are the negative finding reports. So, where perhaps the findings are inconclusive and --

CO-CHAIR TIRSCHWELL: Well, it says you are including people that had a symptom suggestive of a stroke. How do you identify those patients?

DR. SEIDENWURM: Perhaps I think I understand the spirit of your question. What you are concerned about is that this casts too wide a net perhaps in the scans that are included in the measure. I think that that perhaps is a valid criticism, but I think one prefers to cast too wide of a net than a net that is not wide enough in this setting.

CO-CHAIR TIRSCHWELL: Bill?

MEMBER BARSAN: Yes, you spoke to something about the timeliness of the study, so it would be useful to clinicians in places

where they don't have house staff or stroke teams, or whatever. There is nothing in this measure that says anything about timeliness. I mean, it counts just as much if it was reported within 30 minutes as within 30 days.

CO-CHAIR TIRSCHWELL: Risha?

MEMBER GIDWANI: Well, I have sort of the same concern as Bill. The timeframe that is noted here is that the CT or MRI report is within 24 hours of hospital arrival. So, could the developers speak to how this is reconciled with their desire to have timely treatment for stroke?

DR. SEIDENWURM: Yes. I think the intent here was to capture the first one. We don't have always a very good ability to capture the timing of some of these things. There was also a problem of crossing from day to day and some things like that that we were addressing.

CO-CHAIR TIRSCHWELL: Gail?

MEMBER COONEY: I am just confused as to

whether you are measuring whether that report is there at the time of discharge or whether you have got a way of knowing when that report hit the clinician.

DR. SEIDENWURM: This performance measure does not address the issue of timing except for the 24 hours to try to get at the first scan.

You know, we are taking this one step at a time. We would like to address that issue at some point, but at this exact moment we are addressing the content of the report, and timing is a separate topic for another day perhaps.

CO-CHAIR TIRSCHWELL: Dan? And then, Ramon.

MEMBER LABOVITZ: I think the radiology report is a very poor place to try to look for a diagnosis. As it is constructed now, this measure is asking my radiologists to make the diagnosis of TIA or not. They are not good at that.

It also has a problem in that it is looking for intracranial hemorrhage. So, that includes subdural hematomas, traumatic hemorrhages, subarachnoid. What are we really going for here?

CO-CHAIR TIRSCHWELL: Do you want to respond?

DR. SEIDENWURM: Well, each of the hemorrhagic conditions that you named are equally contraindications for the administration of t-PA, I think. So, I think that would be the reason to be inclusive.

CO-CHAIR TIRSCHWELL: Wait. Ramon?

MEMBER RICHMOND: They are also accounting for non-stroke presentations like MS exacerbations, for example, or Todd's paralysis. I mean, this blows your denominator beyond recognition in a way it makes no sense.

CO-CHAIR TIRSCHWELL: Okay. Greg? And then, Michael.

MEMBER KAPINOS: Yes. So, it sounds

like you tried to implicate more the radiologist into the hyper-acute treatment. But 24 hours, as we said, that is not adequate. So, if the real goal is to get radiologists to do a report for t-PA, then it should be within three hours. You should be at the console doing the report, the same stroke neurologists are reading on time, and the majority of radiologists are not doing that at all, even at the very good centers. So, I think it is useless to incorporate this measure.

CO-CHAIR KNOWLTON: David was first, actually. Sorry.

CO-CHAIR TIRSCHWELL: Okay. I think the problem we are struggling with is that this measure is really defined on the wording that is used in the report essentially, what findings are declared to be present or absent. It is taking any report, the report of the first study done on such a patient, if it was done with 24 hours, and people have pointed

out that you don't really -- I mean, that is a heterogenous group that includes people who are in a treatment window which you care about and people who are far outside one that for this it doesn't seem to make much difference.

But my basic concern is it assumes that even the interpretation of their symptoms is static. It assumed by going through and looking at their charts, you get an idea of, you decide which patients presented with the symptoms that would include them. That is not necessarily what their symptoms were perceived to be at the time the study was ordered, which was the basis for their radiological interpretation, so you could get down-scored, because at the time you read the study you didn't know that was what people thought was going on. And the same problems that were already mentioned about figuring out who is in and who is out.

I guess I am just having trouble. It looks like it is an awkward attempt. I think

the idea of trying to make sure you are getting an appropriate interpretation in a timely manner is a good idea, but it looks like some of these problems have led it off the track, and to get back on, I think you would have to narrow the focus to the people who are in a treatment window and come up with practical answers to these other problems about who that is and what you are really measuring, which isn't so much the wording of the report, but the accuracy and timeliness of the information.

CO-CHAIR KNOWLTON: Michael?

MEMBER KAPLITT: So, let me just ask the developer very directly: the numerator statement says "Final reports of the initial CT or MRI that include documentation of the presence or absence" of these things. So, just clarify for me. The CT scan is done within 24 hours. The report comes back a week later and says that there was a documented infarct. That counts in the numerator, is

that correct?

Yes? You are nodding? Yes? Okay.

DR. SEIDENWURM: Yes.

MEMBER KAPLITT: Okay. Tell me how that changes practice or improves stroke care.

DR. SEIDENWURM: Sure. It is rather similar to the argument that was made about coding. There is enormous pressure in the system to have your preliminary report and your final report match with respect to content. So, the data element that we chose was the final report. Because the preliminary report and all these other things appear in a lot of different places, in different charting mechanisms, in different communication schemes, in different institutions. So, in order to have reproducible data elements, we chose that.

Now there are other imperatives in medical care to have the preliminary report and the final report match. So, it is rather analogous to the argument that was made

earlier with respect to coding.

MEMBER KAPLITT: With all due respect, I mean, you are telling me about matching reports, but maybe I am the only one here who is not understanding, but, again, I don't understand how that affects stroke treatment or practice. How a report that is a week later which may be more accurate data, how it is actually affecting practice?

CO-CHAIR TIRSCHWELL: Doesn't it suggest that the report has to be finalized within 24 hours, not a week later?

PARTICIPANT: If I could just clear something up --

CO-CHAIR TIRSCHWELL: Sure.

PARTICIPANT: -- the operand of documenting hemorrhage and mass lesion and acute infarction is the critical part. So, presence or absence, and they have to document all of those three. So, in the final report, sometimes it might say "no hemorrhage," but it doesn't mention mass lesion or acute

infarction. So, language, the presence or absence has to be documented on each of those three elements.

And the final reports are of the initial CT. So, if you have a drip-and-ship, somebody gets CT'ed in a rural area and then they get shipped, there may be another CT done, et cetera. So, the wording is carefully crafted to collect all that information.

MEMBER KAPLITT: Again, I don't understand, though. Okay, so let's take it to the extreme. So, let's say the final report is six months later. Isn't that included in this numerator statement, unless I am misreading something?

DR. SEIDENWURM: Well, of course it would be, but that is a situation that, even in our imperfect system, doesn't happen frequently.

CO-CHAIR TIRSCHWELL: Can I just add one more thing? And I really think we should put it to a vote because then we will see what

happens.

But affecting impact, I think the implication is that I think anybody who has a TIA or a stroke symptom, if they don't get a CT scan, that counts against the facility, is that right, or the physician? It is not just the ones that get the scan? Or is it just the ones that get --

DR. SEIDENWURM: This is just the ones that get the scan.

CO-CHAIR TIRSCHWELL: Okay.

DR. SEIDENWURM: We are requesting that one --

CO-CHAIR TIRSCHWELL: All right

DR. SEIDENWURM: -- narrow issue.

CO-CHAIR TIRSCHWELL: Any last comments? Dave?

MEMBER HACKNEY: I agree with Michael's point. It seems to me that the way the numerator is written, it could be a report generated a year from now. But the limitations on the denominator, to take it a

step further, mean that you are not necessarily looking at the same populations because the denominator has a restriction that has to be either within 24 hours inpatient or in an outpatient setting. But the numerator does not have any such restriction.

MEMBER KAPLITT: No, but the denominator, again, it is only the restriction on when the CT was performed, but the measure is the final report. The denominator places not time restriction on the final report. The restriction is on what the final report is reporting on.

MEMBER HACKNEY: Right. But I am saying that --

MEMBER KAPLITT: The final report can occur anytime.

MEMBER HACKNEY: Right, but it is still not necessarily they don't match. They are not necessarily the same population.

CO-CHAIR TIRSCHWELL: Yes, and I think, you know, if the spirit of the measure was to

urge comprehensively-worded, timely reports, it doesn't seem like this is what -- this isn't what this measure is forcing people to do.

So, I would suggest that we call for a vote on whether there is high impact. Could we do the voting, please?

(Vote taken.)

MS. THEBERGE: One high, 1 moderate, 15 low, 5 insufficient evidence.

CO-CHAIR TIRSCHWELL: So, I think that concludes our work on this measure for now.

DR. BURSTIN: I just want to let the group know that there was a measure endorsed about a year ago that was head CT or MRI scan results for acute ischemic stroke or hemorrhagic stroke with an interpretation within 45 minutes of arrival. So, that has already been endorsed.

MEMBER KAPLITT: That is useful.

CO-CHAIR TIRSCHWELL: Can we go back to 0439?

CO-CHAIR KNOWLTON: 0439 is "Discharged on Statin Medication". Jolynn, I believe you are doing this.

CO-CHAIR TIRSCHWELL: Salina wants to make a comment.

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

MEMBER WADDY: So, if that was endorsed, if that measure was endorsed a year ago, I mean, will that measure continue going forward? Okay.

DR. BURSTIN: It was a measure preferred by CMS. I don't know if they have put it forward for any uses yet. I think there may still be testing.

CO-CHAIR KNOWLTON: Jolynn?

MEMBER SUKO: Yes, this is discharged on statin medication. This is a measure that was originally endorsed on July 31st, 2008. It includes ischemic stroke patients with LDL greater than or equal to 100 who are on a lipid-lowering medication prior to arrival, who are prescribed statin medication at

hospital discharge. And the denominator is ischemic patients with an LDL greater to or equal to 100 or LDL not measured or who were on lipid-lowering medication prior to hospital arrival.

The Work Group, on impact, there was general agreement important to measure. Again, this is a similar theme to some of our others discharged on antithrombotics in the sense that we have a very high performance rate around this. But, again, the Work Group, six rated it high, one medium.

CO-CHAIR KNOWLTON: Questions or comments?

(No response.)

Seeing none, let's vote.

(Vote taken.)

MS. THEBERGE: We need one more vote. Or somebody stepped away.

Nineteen high, 1 moderate, 1 low.

CO-CHAIR KNOWLTON: Okay. Jolynn, evidence?

MEMBER SUKO: Evidence, yes, this was unanimously seen as having high evidence behind it by the Work Group, not much further discussion.

CO-CHAIR KNOWLTON: Any discussion here?

MEMBER J. BAUTISTA: Well, don't the Work Group scores reflect medium for quality or moderate for quality?

MEMBER SUKO: Oh, sorry.

CO-CHAIR KNOWLTON: I don't know who is talking. I can't hear --

MEMBER KAPINOS: So, can somebody explain again the criteria for how many RCTs do we need to vote for?

DR. BURSTIN: It is also in your Quick Guide at your table, if you would like. So, it specifically says for quantity, for high, it is five-plus; moderate, two to four; low, one. But you also have this sheet in front of you.

CO-CHAIR KNOWLTON: Any comments on evidence?

CO-CHAIR TIRSCHWELL: You know, I guess I would comment that, although the SPARCL study is the dominant theme here, there are other large statin trials that had a subset that had already had a stroke who also seemed to benefit from it. So, I think, in my opinion, there is more than one study that shows the benefit here.

CO-CHAIR KNOWLTON: Anybody else?

(No response.)

Okay. Let's vote.

(Vote taken.)

MS. THEBERGE: We are still short one response.

All right, can everyone vote one more time?

Somebody stepped away?

Nineteen yes, 2 no.

CO-CHAIR KNOWLTON: Okay. Performance gap.

MEMBER SUKO: Like in our discussion around discharge on antithrombotics, this

performance is very, very high for this measure. There is some evidence that there are disparities in minority populations.

The Committee rated this four high and three medium.

CO-CHAIR KNOWLTON: Comments?

(No response.)

Okay. Let's vote.

(Vote taken.)

MS. THEBERGE: Ten high, 10 moderate, 2 low.

CO-CHAIR KNOWLTON: Okay. Scientific acceptability, starting with reliability.

MEMBER SUKO: Reliability, six rates this as high, one medium. And I think the medium rating was based largely on presentation of the data.

CO-CHAIR KNOWLTON: Comments? Questions?

(No response.)

Let's vote.

(Vote taken.)

MS. THEBERGE: Nineteen high, 3 moderate.

CO-CHAIR KNOWLTON: Validity?

MEMBER SUKO: A similar theme for reliability (sic), six voted high, one voted moderate.

CO-CHAIR KNOWLTON: This is validity, right?

MEMBER SUKO: Yes.

CO-CHAIR KNOWLTON: Okay. Questions or comments?

(No response.)

Okay. Let's vote.

Oh, I'm sorry. Salina?

MEMBER WADDY: I do agree with the comment that was made by the group regarding which patients really we should be talking about. I mean, there is some evidence that globally stroke patients may benefit, but they did make the comment regarding atrial fibrillation, cardioembolic stroke, and the subarachnoid hemorrhage patients.

Is that going to be teased out in some way or is it just globally all stroke patients who should receive the statin?

CO-CHAIR TIRSCHWELL: Well, in the numerator it says ischemic stroke. So, I think that removes the subarachnoid hemorrhage patients from the group pretty rapidly.

And then, you know, my own personal opinion about the atrial fibrillation is that, if you meet the other criteria for the measure, she should probably be on a statin anyway, although --

MEMBER WADDY: No, I agree with that as well. But, in terms of the evidence that is really out there, I mean, that is a bit of a more controversial topic.

MEMBER BARSAN: There are also dissections that fit into there, too, and stop. So, if you had a dissection and you don't measure cholesterol, you get dinged if you don't start them on a statin.

DR. SCHWAMM: David, can I just clarify?

CO-CHAIR TIRSCHWELL: Sure.

DR. SCHWAMM: So, in the measure logic, patients who have a documented reason for not having an indication for statins are eliminated from the denominator. So, if you wrote in the record, "quad-dissection, no atherosclerosis, no indication for statin," you would be excluded from the measure.

So, the only way you would get, quote/unquote, "dinged" is if you didn't document any reason why statin wasn't considered and simply discharged the patient.

CO-CHAIR TIRSCHWELL: Anything else?

MEMBER KAPINOS: We will have to really like create some templates. Once those measures are implemented, yes, there will be a great need for templates for clinicians to be matching all those like sets of -- it becomes so convoluted not to just follow your pathway as a clinician, but just to match all those like --

CO-CHAIR TIRSCHWELL: So, Greg, I am not

sure where you are practicing, but these are already implemented, and the templates already exist in most cases. So, you are absolutely right, and it gets to, are we doing what is right for the patient? We then also have to document that we are doing what is right for the patient. That is just an inherent burden in this whole quality measure phenomena.

MEMBER KAPINOS: Oh, patient did not have an MI. That is why I didn't give him -- I mean, did not have MS and did not have --

CO-CHAIR TIRSCHWELL: Don't give them any ideas.

CO-CHAIR KNOWLTON: Other comments? Salina?

MEMBER WADDY: I mean, to me, it seems simpler to at least list the ones that have the strongest evidence. If they have intracranial stenosis, extracranial stenosis, rather than parsing out which people may think lacunar strokes should or should not -- it just gets to be very --

CO-CHAIR TIRSCHWELL: So, I guess if you are getting to ischemic stroke subtypes, other than the a-fib, which is pretty clean, whether somebody else has an atherothrombotic ischemic stroke or not, and then relying -- in many cases, it is some sort of person in the QI department who is trying to figure this out. I think it is a real setup for problems.

CO-CHAIR KNOWLTON: Okay. Anything else?

(No response.)

On some of these things, the issues are interesting, but they don't speak to the measure that is before us. You know, they are suggestions that we might make for later.

Other thoughts on this?

(No response.)

Then, let's vote. This is on validity.

(Vote taken.)

MS. THEBERGE: We need one more vote.

Twelve high, 9 moderate, 1 low.

CO-CHAIR KNOWLTON: Okay. Usability.

MEMBER SUKO: So the subgroup unanimously ranked this as high. And I would note that this has been in practice and undergone cycles of improvement for at least three years.

CO-CHAIR KNOWLTON: Burning issues?

(No response.)

Okay. Let's vote.

(Vote taken.)

MS. THEBERGE: Twenty-one high, 1 moderate.

CO-CHAIR KNOWLTON: Feasibility.

MEMBER SUKO: Again, similar to usability, this has been in place for quite some time. It was ranked very high by the subgroup.

CO-CHAIR KNOWLTON: Comments?

(No response.)

Okay. Let's vote.

(Vote taken.)

MS. THEBERGE: One more vote.

Twenty high, 1 moderate, 1 low.

CO-CHAIR KNOWLTON: And overall suitability for endorsement.

MEMBER SUKO: So, the subgroup unanimously recommended this for endorsement.

CO-CHAIR KNOWLTON: Comments?

(No response.)

Okay. Let's go.

(Vote taken.)

MS. THEBERGE: All right. Twenty. Still at 21. Okay.

Twenty-one yes, 1 no.

CO-CHAIR KNOWLTON: Okay. According to my agenda, we should now have lunch.

(Laughter.)

Just saying. Just saying.

CO-CHAIR TIRSCHWELL: All right. We are going to plug right along, and I guess now we are on Stroke Measure No. 1 from The Joint Commission, Venous Thromboembolism.

And, Jocelyn, are you going to start us off?

MEMBER J. BAUTISTA: Okay. So, this

measure is titled, "Venous Thromboembolism Prophylaxis". And it is meant to capture those patients with ischemic or hemorrhagic stroke who either receive VTE prophylaxis or have documentation of why prophylaxis was not given by the second hospital day.

So, the numerator is as mentioned, and the denominator is all ischemic or hemorrhagic stroke patients 18 years of age and older, between a length of stay of 2 and 120 days, and they have exclusions as listed.

And they define receive VTE prophylaxis as either pharmacological or mechanical. And under mechanical, they include IPCs, intermittent pneumatic compression devices, as well as venous foot pumps. But they did not include the TED hose compression hose.

And in order to meet the measure, you simply have to document that it was given for the first time by the second hospital day.

And also to meet the measure, if there is documentation of why not VTE prophylaxis

was given by a licensed provider, that is also considered a pass.

Patient refusal is acceptable as a reason for not giving.

CO-CHAIR TIRSCHWELL: Do you want to comment on impact?

MEMBER J. BAUTISTA: Okay. So, impact. I think the Work Group, five rated it as high, one moderate. Stroke, I believe per the measure submission document, PE is seen in about 1 percent of stroke patients, DVT in a higher percentage, and VTE prophylaxis reduces PE by 50 percent, 50 to 70 percent.

CO-CHAIR TIRSCHWELL: Anybody want to comment on impact?

(No response.)

Okay. Let's go ahead and vote.

Wait. Wait for the timer. Okay, now vote.

(Vote taken.)

MS. THEBERGE: Seventeen high, 5 moderate.

CO-CHAIR TIRSCHWELL: Okay. Moving on to evidence -- Greg, go ahead.

MEMBER KAPINOS: Why did we lump together the pharmacological and the mechanical devices? Because, actually, for the timeliness issue, we could actually separate those two and say that there shouldn't be any reason not to apply at least the mechanical devices on day one, and then, actually, the pharmacological one on day two.

So, by lumping it together, when we know that there is definitely more risk of hemorrhagic complication with the pharmacological one, I am scared that, then, you could become compliant by just, oh, now it is day two after a stroke; we need to think about applying those compression devices and starting the hep subQ. To me, that is not the best practice. It is to start with the compression devices and, then, think about the pharmacological one.

CO-CHAIR TIRSCHWELL: Risha?

MEMBER GIDWANI: In reviewing the evidence, it looks as though, based off of the data the developers presented, that the actual mechanical prophylaxis was not found to reduce VTE, but the low-molecular-weight heparin was. So, why are those not be differentiated?

CO-CHAIR TIRSCHWELL: Does the developer want to comment on that?

DR. SCHWAMM: Sure. I will take the first comment.

I just want to call your attention to the fact that this measure is actually for all stroke patients, not just ischemic stroke patients, in line with the previous comments about having the largest applicable population.

And for many patients with hemorrhagic stroke, low-molecular-weight heparin or other anticoagulants are contraindicated, at least in the first several days after the event.

And so, the effort here is to ensure that appropriate therapy is being initiated in

all patients by the second hospital day. So, I think that is the emphasis here. Rather than just segregating out a smaller denominator who would be eligible for chemical prophylaxis, the decision was made to incorporate a larger denominator.

CO-CHAIR TIRSCHWELL: And I would just comment that the way the days are coded, it is possible that you could be admitted at 11:59 p.m., and you would be on your second hospital day two minutes later. And so, we thought we should give them at least two minutes to get the DVT prophylaxis going.

(Laughter.)

So, I think that is why it is not a one-day thing. If there was a way to accurately account for within 24 hours, maybe that would be preferable, but it doesn't seem like that is easily implemented.

Risha, go ahead.

MEMBER GIDWANI: I am sorry. I am not a clinician. So, maybe that is where my lack

of understanding stems from. But when I see a sentence that says that "Graduated compression stockings were not found to reduce VTE risk or death within the first seven days post-stroke," I am wondering, and then, in light of the comment that you just made, I am wondering, is it potentially, then, more reasonable to talk about doing chemoprophylaxis in a smaller group of patients?

DR. SCHWAMM: So, maybe just to clarify, graduated compression stockings, also called TED hose, are not sufficient for DVT prophylaxis. Pneumatic compression devices or sequential compression devices -- pneumoboots is what they are often called -- are considered appropriate. So, actually, the stockings alone do not give you credit for this measure, even though in the broader hospital-based VTE prophylaxis measure, which includes all sorts of patients at very low risk for VTE, for those patients, stockings

alone would be sufficient. So, I think that is the piece that distinguishes this.

CO-CHAIR TIRSCHWELL: Any other comments about evidence? Questions?

MEMBER J. BAUTISTA: Yes.

CO-CHAIR TIRSCHWELL: Jocelyn?

MEMBER J. BAUTISTA: The measure includes venous foot pumps as a valid prophylaxis. And yet, I didn't see that in the evidence, that venous foot pumps had been demonstrated --

CO-CHAIR TIRSCHWELL: I have been practicing for 20 years. I don't even know what a venous foot pump is.

(Laughter.)

What is a venous foot pump?

DR. SCHWAMM: So, I think you see them largely in places like the burn unit. They are essentially intermittent compression devices that don't require application to the entire limb. I think their use is quite rare. But I think they are FDA-approved for use in

prevention of VTE prophylaxis, which is probably why they are still listed as acceptable devices.

CO-CHAIR TIRSCHWELL: Okay. Thank you for that clarification.

Any other comments? Yes, Gwen?

MEMBER BUHR: So, the evidence presented here does say that, going back to what Risha was saying, that intermittent compression devices, which is not the TED hose, where it says here "with a non-significant trend toward decreased risk of VTE," but that is non-significant, with no impact on deaths. So, I am still questioning the quality of the evidence here.

CO-CHAIR TIRSCHWELL: I mean, my impression is that the evidence is not as strong for SCDs. I don't have much more to add to that.

Dr. Schwamm, do you?

DR. SCHWAMM: The real question, I think, is in patients who are not eligible for

chemoprophylaxis, these appear to be better than no form of treatment.

So, I think you are right. I think the evidence for those is somewhat weaker, but I think that it is still important for patients to go with some form of treatment to reduce the risk of DVT/VTE if they can't be treated with chemoprophylaxis.

CO-CHAIR TIRSCHWELL: Okay. Let's go ahead and vote, then, please.

(Vote taken.)

MS. THEBERGE: Fifteen yes, 7 no.

CO-CHAIR TIRSCHWELL: Okay. Going back to 1b, evidence of a performance gap.

MEMBER J. BAUTISTA: So, the developers note performance of 93 percent for third quarter 2011. They also report another rate of 88 percent for the same time period. And from 2003, the rate increased from 74 to 90 percent. So, the Work Group did feel there was a gap.

CO-CHAIR TIRSCHWELL: Comments?

Questions?

(No response.)

No? Let's go ahead and -- oh, go ahead, Greg.

MEMBER KAPINOS: The gap, if we broke it into a subgroup within pharmacological versus mechanical, it would have been interesting, also, to see.

CO-CHAIR TIRSCHWELL: So, there is a suggestion that pharmacological versus chemical (sic) be broken down. For that matter, it would be interesting to look at ischemic versus hemorrhagic as well.

Let's go ahead and vote on whether there is a performance gap.

(Vote taken.)

MS. THEBERGE: Eleven high, 9 moderate, 2 low.

CO-CHAIR TIRSCHWELL: Okay. Very good. So, I think that means we are moving on to Question 2, scientific acceptability, and starting with reliability.

MEMBER J. BAUTISTA: So, reliability was tested at the data-element level only. Otherwise, I don't think there were any specific concerns.

CO-CHAIR TIRSCHWELL: Okay. Any questions or comments on reliability?

(No response.)

Let's go ahead and vote.

(Vote taken.)

MS. THEBERGE: Fourteen high and 8 moderate.

CO-CHAIR TIRSCHWELL: Validity, 2b.

MEMBER J. BAUTISTA: The Work Group thought, well, half of us thought it was high, the other half moderate validity. Some questioned the exclusions in the measure, which we have already discussed.

CO-CHAIR TIRSCHWELL: Okay. So, the question of the exclusions might reflect this vote.

Any comments or questions?

(No response.)

Let's go ahead and vote.

(Vote taken.)

MS. THEBERGE: Five high, 16 moderate, 1 low.

CO-CHAIR TIRSCHWELL: All right. Moving on to usability.

MEMBER J. BAUTISTA: Yes. So, this measure has been in use. I don't think we had any issues with usability.

CO-CHAIR TIRSCHWELL: Comments? Questions?

(No response.)

Vote.

(Vote taken.)

MS. THEBERGE: Fifteen high, 7 moderate.

CO-CHAIR TIRSCHWELL: Okay, and the second-to-last, feasibility.

MEMBER J. BAUTISTA: Yes, similar to most of the other measures, there is a fair amount of abstraction that is required from this measure, but it has been in use for several years now.

CO-CHAIR TIRSCHWELL: Let's go ahead and vote if there are no comments.

(Vote taken.)

MS. THEBERGE: We have 21 votes.

CO-CHAIR TIRSCHWELL: Whoever voted early, vote again.

Everybody vote one more time, please. Sorry.

MS. THEBERGE: Nine high, 12 moderate.

CO-CHAIR TIRSCHWELL: And then, finally, overall suitability for endorsement.

Any further comments, Jocelyn?

MEMBER J. BAUTISTA: No, and we will get to the fact that it overlaps with another measure later, but the group did think it was suitable as stands.

CO-CHAIR TIRSCHWELL: Okay. Wait until you see the timer before you vote. Okay, go ahead and vote now.

(Vote taken.)

There we go.

MS. THEBERGE: Twenty-one yes, 1 no.

MEMBER KAPINOS: I just realized that, for the first time in those measures, somebody mentioned we have an exclusion about patient refusal. How come patient refusal didn't make it through all the other measures, like patient can refuse t-PA and he can refuse everything?

CO-CHAIR TIRSCHWELL: My guess is that many of the other measures that say physician-documented reason for exclusion, you would be able to probably apply that.

MEMBER KAPINOS: But why does it stand out in this measure?

MEMBER J. BAUTISTA: Because I think it is a real issue with the devices. Patients do refuse the devices because they are uncomfortable.

MEMBER KAPINOS: Oh, yes. Okay.

CO-CHAIR TIRSCHWELL: Or the shots in the stomach, for that matter.

Dr. Schwamm, did you have a final comment?

DR. SCHWAMM: No.

CO-CHAIR TIRSCHWELL: No? Okay.

CO-CHAIR KNOWLTON: Let's move on to 0240, Deep Venous Thrombosis Prophylaxis for Ischemic Stroke or Intracranial Hemorrhage.

This is maintenance on an original endorsement, and let's see where we are.

Jack? You are presenting it?

MEMBER SCARIANO: No.

CO-CHAIR KNOWLTON: Oh, didn't you present it in the Work Group?

Who else is in that Work Group that remembers it?

CO-CHAIR TIRSCHWELL: Does anybody remember presenting Measure 0240?

CO-CHAIR KNOWLTON: Salina, do you remember that?

MEMBER J. BAUTISTA: I remember that Jack was assigned, but he didn't realize he was assigned.

CO-CHAIR KNOWLTON: Okay. Well, does anybody want to pick it up or shall we just go

through it?

(No response.)

Okay. This is a measure that was originally endorsed in May of 2007. It is the percentage of patients aged 18 and older with a diagnosis of ischemic stroke or intracranial hemorrhage who are administered DVT prophylaxis by the end of hospital day two.

The numerator and denominator statements appear to be straightforward. And the Working Group -- I was not on this Working Group -- but the Working Group felt that it had impact: four high, one medium.

Anybody want to comment on this? Yes?

MEMBER J. BAUTISTA: So, I am in the Work Group. It is very similar to the measure I just presented. So, I could kind of walk through it.

This is something I didn't realize when we were discussing it in the Work Group, but patients who have expired during the inpatient stay are excluded from the measure. So, you

know, potentially, you have a patient who comes in with an ischemic stroke, develops a DVT/PE, dies a week later, and they are excluded from the measure.

CO-CHAIR KNOWLTON: Does this come under bury your mistakes?

(Laughter.)

Does the developer want to comment on this wrinkle?

MEMBER KAPINOS: The other thing is the VTE prophylaxis is a better terminology than DVT because what you are proven to do is not only to prevent DVTs, but PEs. So, actually, the other measure is using a better --

CO-CHAIR KNOWLTON: Yes, but let's not harmonize them yet. Let's just go through. It will take you a longer time to do that. We will harmonize.

Did the developer demure on the --

MS. HANLEY: I mean, I think the exclusion criteria to not include patients who expired is simply more of a methodological one

to make sure that we have a more clear-cut denominator.

MEMBER J. BAUTISTA: It would make more sense to me to exclude patients who have expired by day two, obviously, but why exclude the others?

CO-CHAIR TIRSCHWELL: I am guessing to some degree, but my thought is that a lot of the patients that expire after a stroke, it is often because it was very serious. And I think that is probably what you are referring by saying that they expired by day two. I guess I don't know that there is a guarantee that a severe stroke patient would die by day two.

MEMBER J. BAUTISTA: No, I mentioned day two because this is a measure that is measured at day two.

CO-CHAIR KNOWLTON: Well, we can vote on whether that impacts the impact statement or whether we want to make it a recommendation to clean up the measure to the measure developer.

MEMBER KAPLITT: But it is a good point relative to the impact, because the developer says that part of the impact of this is that 10 percent, they say, of stroke deaths are due to PEs, right, which is a significant percentage.

CO-CHAIR KNOWLTON: Yes.

MEMBER KAPLITT: If you are going to be excluding those patients, it relates to impact. I am just reading from what it says here.

CO-CHAIR KNOWLTON: This is also, as I read it -- this is currently being measured, is that correct? Yes? Okay.

Other comments on impact for a concern about the expiration of the patient?

(No response.)

We can't edit the measures.

So, anything else?

(No response.)

Okay. Let's vote on impact.

(Vote taken.)

MS. THEBERGE: We are still at 21 responses.

I think the problem may be with the right side of the table. There might be too much stuff in between you and the receiver.

Okay. Seven high, 11 moderate, 3 low, 1 insufficient evidence.

CO-CHAIR KNOWLTON: And we move on.

Jocelyn, will you continue to go through this?

MEMBER J. BAUTISTA: Sure, sure.

CO-CHAIR KNOWLTON: Thank you.

MEMBER J. BAUTISTA: So, for the evidence, I think similar issues, pharmacologic prophylaxis is strong evidence, less so for the mechanical devices. But the group overall thought there was sufficient evidence.

CO-CHAIR KNOWLTON: Any questions or comments?

(No response.)

I am going to move through this quickly

because it is the same type of issues, unless something emerges that is different.

Anything else?

(No response.)

All right, let's vote.

(Vote taken.)

Okay, got that.

MS. THEBERGE: Seventeen yes, 5 no.

CO-CHAIR KNOWLTON: Okay. Performance gap.

MEMBER J. BAUTISTA: Right. So, the estimated performance -- current performance is 79 to 84 percent, based on 2009-2010 data.

CO-CHAIR KNOWLTON: Questions?

(No response.)

Vote.

(Vote taken.)

You right-wingers want to aim carefully over there.

MS. THEBERGE: We are still at 21.

There we go.

Fourteen high, 8 moderate.

CO-CHAIR KNOWLTON: Scientific acceptability, reliability.

MEMBER J. BAUTISTA: All right. So, two of us voted high, two moderate. No specific comments.

CO-CHAIR KNOWLTON: Any comments?

(No response.)

Okay. Let's vote.

(Vote taken.)

MS. THEBERGE: Nine high, 13 moderate.

CO-CHAIR KNOWLTON: Okay. Validity, I believe. Yes, validity.

MEMBER J. BAUTISTA: There was some discussion about exclusions and reasons, but similar to the past discussion.

CO-CHAIR KNOWLTON: Comments?

(No response.)

Vote.

(Vote taken.)

MS. THEBERGE: Five high, 16 moderate, 1 low.

CO-CHAIR KNOWLTON: Okay. Usability.

MEMBER J. BAUTISTA: So, the measure is in use currently.

CO-CHAIR KNOWLTON: Questions?

MEMBER J. BAUTISTA: No real issues.

CO-CHAIR KNOWLTON: Comments?

(No response.)

Okay.

(Vote taken.)

MS. THEBERGE: Fifteen high, 7 moderate.

CO-CHAIR KNOWLTON: Okay. Feasibility.

MEMBER J. BAUTISTA: Again, there are data-abstraction issues, but no significant concerns.

CO-CHAIR KNOWLTON: Okay. Let's vote.

(Vote taken.)

MS. THEBERGE: Twelve high, 10 moderate.

CO-CHAIR KNOWLTON: And suitability for endorsement. There is nothing really to say about it. Let's just vote.

(Vote taken.)

MS. THEBERGE: Twenty-one high -- I'm sorry -- 21 yes, 1 no.

CO-CHAIR TIRSCHWELL: Moving along, the next measure on the docket is Stroke Measure No. 3 from The Joint Commission, Anticoagulation Therapy for A-Fib/Flutter.

Fred, do you want to start us off?

MEMBER TOLIN: Well, thank you.

This is a measure that is already in place. It has been for several years. And as pointed out, it is one of the eight measures used by The Joint Commission.

The need for anticoagulation in patients who have non-valvular atrial fibrillation and flutter, the medical evidence is really not very controversial, and the exclusions are relatively straightforward.

The only real thing that came up in this was a question of timing in terms of how soon after stroke should the antithrombotic therapy or anticoagulant therapy be initiated. As pointed out on one of the other measures, this is also silent as to the agent used. Traditionally, warfarin, as I think everybody

is aware, was the only agent available, but now there are newer agents available. So, it doesn't define which agent is used, just the fact that an anticoagulant is ordered.

CO-CHAIR TIRSCHWELL: Right. So, to clarify, aspirin wouldn't count.

MEMBER TOLIN: Aspirin would not count. Warfarin or --

CO-CHAIR TIRSCHWELL: But the newer agents and warfarin --

MEMBER TOLIN: One of the newer drugs would.

CO-CHAIR TIRSCHWELL: Yes. And to me, that seems okay.

So, any other questions or comments about impact before we vote on impact?

(No response.)

Let's go ahead and activate the system and now vote.

(Vote taken.)

MS. THEBERGE: Twenty-one high, 1 moderate.

CO-CHAIR TIRSCHWELL: Thank you.

Moving on to 1c, evidence.

MEMBER TOLIN: In the Committee meeting, the vote was three for yes, one for no.

The evidence is certainly available and, as I pointed out previously, is not controversial. The only question, again, came up in terms of timing and, also, the agent, the lack of specificity of the agent.

CO-CHAIR TIRSCHWELL: Okay. Ramon, you had a comment?

MEMBER R. BAUTISTA: Yes. Just for the stroke neurologists here, is this true, 52 percent of patients do not meet the measure? Is that correct? So, half the patients with atrial fib do not receive anticoagulation? I mean, I don't know the answer.

CO-CHAIR TIRSCHWELL: I think those are some more community-based --

MEMBER R. BAUTISTA: That is true, though, huh?

CO-CHAIR TIRSCHWELL: -- things, but

there's lots of literature suggesting a large gap, not necessarily in hospitalized patients at the time of discharge, but in a broader view, I think, of the effective treatment with atrial fibrillation.

And I guess some of those data are older. I am hoping it is not as big a gap these days, but I am not 100 percent sure.

As far as the timing issue goes, I guess that speaks to the evidence a little bit or the validity. I agree with raising the question about timing. I don't know that there is -- certainly, long-term, secondary prevention, there is no doubt about the anticoagulation.

There was one particular study that showed that in the first two weeks full-dose low-molecular-weight heparin, so anticoagulation, versus just a regular aspirin a day, there was no difference in outcomes at the two-week mark. And in fact, the low-molecular-weight heparin group had more

complications. And then, of course, everybody was anticoagulated long-term.

So, I often suggest, especially with larger strokes, that we don't start until two weeks after they are discharged. And we are not getting dinged. So, somehow, maybe that is written down as a reason why they are excluded, but I do think it is an issue.

Salina?

MEMBER WADDY: Actually, there are a couple of other reasons why people don't prescribe or may not prescribe. One is in minority populations there has been a disparity identified, both in the prescription as well as in patients who, for whatever reason, have been told that they had atrial fibrillation and they don't take the medication.

But the other major issue is people who have significant deficits who either (a) don't have a family member that can appropriately, a family member or a caretaker to administer

coumadin, as well as people who are going on to rehabilitation. There is the concern in some of the physicians that these people may be at fall risks. And so, they will just put them on aspirin instead in the acute period.

CO-CHAIR TIRSCHWELL: In the acute period, and those are probably deemed to be valid reasons, as long as you, of course, document that you have done so.

So, any other comments on evidence?

(No response.)

I suggest we activate the voting and go ahead and vote on evidence.

(Vote taken.)

MS. THEBERGE: Twenty-one yes, 1 no.

CO-CHAIR TIRSCHWELL: Okay. Great.

And then, back to 1b, performance gap.

MEMBER TOLIN: So, in the time period since 2008, when this measure was first initiated or enacted, the actual performance has increased. And so, the performance gap is less than it was previously, and the number we

are looking at now is probably somewhere around 95 percent. So, there is not a significant performance gap, which is similar to some of the other measures we have talked about earlier today.

CO-CHAIR TIRSCHWELL: Any comments? Questions about that?

MEMBER TOLIN: The only other comment, I think it has also been pointed out there is a gap, a bigger gap, in some minority populations. I should point that out also.

CO-CHAIR TIRSCHWELL: And there is evidence for that in the literature, although not necessarily as the data presented from the measure itself, from The Joint Commission. They don't have those data. I don't know; Get With The Guidelines probably does, but they are not 100 percent overlap there.

MEMBER WADDY: Right. Some of them don't tease out the minority.

CO-CHAIR TIRSCHWELL: Yes. Anyway, performance gap, let's go ahead and vote.

(Vote taken.)

MS. THEBERGE: Eight high, 12 moderate, 2 low.

CO-CHAIR TIRSCHWELL: So, we proceed along to 2a, which is reliability.

MEMBER TOLIN: In Committee, this was all unanimously high in terms of -- I'm sorry -- moderate in terms of reliability, not high.

Thank you, Karen. She is giving me that evil eye.

(Laughter.)

One of the concerns was that some of the exceptions are not exactly well-spelled-out, and there is some inconsistent evidence in the exceptions in the denominator.

CO-CHAIR TIRSCHWELL: Does anybody want to comment on the lack of consistency with the evidence?

(No response.)

No?

The comment from the Work Group call, that certain elements only had an 82 or 85

percent agreement.

MEMBER J. BAUTISTA: The developer commented on that issue. If I remember correctly, the issue was that some of the newer anticoagulants weren't yet on the list, and that caused some confusion among the abstractors.

CO-CHAIR TIRSCHWELL: I see. Okay. Fair enough.

Any other concerns? Comments?

(No response.)

Let's go ahead and vote on reliability.

(Vote taken.)

MS. THEBERGE: Four high, 16 moderate, 2 low.

CO-CHAIR TIRSCHWELL: Okay. Next is validity.

MEMBER TOLIN: And similar statements made about validity.

CO-CHAIR TIRSCHWELL: Comments or questions?

(No response.)

Okay. Let's activate the system and now vote.

(Vote taken.)

MS. THEBERGE: Eight high, 13 moderate, 1 low.

CO-CHAIR TIRSCHWELL: Okay. Continue to proceed along. We are up to usability.

MEMBER TOLIN: Also, sort of a split feeling in the Work Group meeting about the usability. As I pointed out earlier, this has been in place for several years. It started out not much lower, and there has been some improvement in the percentage of patients who are receiving anticoagulation therapy. But, again, it is already being used.

CO-CHAIR TIRSCHWELL: Okay. Salina?

MEMBER WADDY: So, I think one of the challenges is not only the treatment of atrial fibrillation. So, this is more of a note to the NQF. It is the actual detection of atrial fibrillation. There is a lot of variability across the country in terms of whether or not

people place patients on telemetry, whether or not they just do a quick screening, EKG.

And so, you are going to have a huge discrepancy in terms of the number of patients that are actually detected. As well, there was up until about a year ago the assumption that, for example, African-Americans didn't have atrial fibrillation until the ERIC data came out. And so, patients weren't even being screened. There were whole hospital systems that didn't have telemetry access, particularly in underserved communities, as well as in some rural and remote places.

So, both of the measures, that I am saying I don't think there is a third measure, is really just on the treatment with anticoagulation, but it really doesn't address the underdiagnosis.

CO-CHAIR TIRSCHWELL: I think that is a great point.

DR. WINKLER: Just as a comment, these are the kinds of suggestions that we collect

from you, because it sounds like there's perhaps another performance measure around screening and diagnosis and identifying appropriate patients.

And we do create a recommendation for measure development kind of section that goes along with all your other recommendations. So, I think we can put it in there.

CO-CHAIR TIRSCHWELL: Fantastic.

Any other comments on usability?

(No response.)

Let's go ahead and -- oh, yes? Go ahead.

MEMBER BARRETT: In the same vein, with regard to public reporting, the issue that was brought up before about rehabilitation and falls and identification of medication adherence is an area for future development, particularly because future medications will not be able to be monitored objectively as well.

CO-CHAIR TIRSCHWELL: Okay. Thank you.

Any more comments?

(No response.)

Let's go ahead and vote on usability.

(Vote taken.)

MS. THEBERGE: We are at 20.

Thirteen high, 8 moderate, 1 low.

CO-CHAIR TIRSCHWELL: Then, feasibility.

MEMBER TOLIN: Again, this is something that is already being used. As Salina pointed out, probably there is some underdiagnosis of some of these things, particularly underlying atrial fibrillation.

CO-CHAIR TIRSCHWELL: Any comments on feasibility?

(No response.)

Let's go ahead and vote.

(Vote taken.)

MS. THEBERGE: Thirteen high, 9 moderate.

CO-CHAIR TIRSCHWELL: And then, finally, overall suitability for endorsement. Anybody want to make some comments?

MEMBER TOLIN: The only thing, again, this is a measure that is already in place, and it is very similar to one of the other measures which I know we will be discussing later, David.

CO-CHAIR TIRSCHWELL: Yes. Let's go ahead and activate the voting. Go ahead and vote.

(Vote taken.)

MS. THEBERGE: Twenty-two yes.

CO-CHAIR TIRSCHWELL: Fabulous.

CO-CHAIR KNOWLTON: Moving on to 0241, Stroke and Stroke Rehab, Anticoagulation Therapy Prescribed by Atrial Fib at Discharge.

This is clinician level, Reva tells me.

It is an original endorsement from May of 2007. So, this is a maintenance action.

And Salina?

MEMBER WADDY: Sure. Yes. Hi.

So, yes, this is already in place and was originally endorsed in 2007. This is different in that the previous atrial

fibrillation treatment question was based on the level of the facility; whereas, this one is for the clinician or group practice.

And it really is just is to detect the number of patients that are actually treated with anticoagulation who either have TIA or ischemic stroke; whereas, the other measure was purely, I believe, ischemic stroke.

Let's see. So, really, there wasn't a whole lot of additional comments beyond that.

This adds the TIA as well as the level of the clinician and the group practice.

In terms of impact --

CO-CHAIR KNOWLTON: Yes.

MEMBER WADDY: -- everyone agreed that this was either high- or moderate-level impact.

CO-CHAIR KNOWLTON: Questions? Comments?

(No response.)

Vote.

Wait for the little prompt. There you

go.

(Vote taken.)

MS. THEBERGE: We are at 18 responses. All right. We are short one.

CO-CHAIR KNOWLTON: David is not here.

MS. THEBERGE: There we go.

Twenty high, 1 moderate.

CO-CHAIR KNOWLTON: Okay. Salina?

MEMBER WADDY: So, in terms of the evidence, the group, there was actually some disagreement for a lot of these questions, but they agreed that there was a significant quantity of information, however, some thought that there was insufficient data. And there was largely agreement that the quality was high or moderate.

We did think -- and this was kind of a recurring theme -- that there were some patients, it wasn't really clear in patients who may have intermediate risk from the atrial fibrillation really what to do with those patients who, for example, only had a stroke,

but didn't have the other elements that would push them over into anticoagulation, particularly if they had a TIA.

CO-CHAIR KNOWLTON: Comments?

(No response.)

Vote.

(Vote taken.)

MS. JOHNSON: We are going to do a redo on that one.

MS. THEBERGE: Sorry about that.

We are at 21.

Twenty yes, 2 no.

MEMBER WADDY: So, in terms of the performance gap, as you can see, this measure was at treatment in only 79 percent of patients, and that is different from the other measure, but this measure, as I mentioned, also includes TIAs as well as it has a different level of analysis. So, it is different, for whatever reason.

CO-CHAIR KNOWLTON: Questions?

(No response.)

Vote.

MEMBER WADDY: And there has been a disparity in care as well in the population.

(Vote taken.)

MS. THEBERGE: Seventeen high, 5 moderate.

CO-CHAIR KNOWLTON: Okay. Scientific acceptability, starting with reliability.

MEMBER WADDY: So, in terms of reliability, everyone thought that this was at least above the low level, but there was a mixed opinion. There were some questions regarding -- or I actually brought up the question in terms of how atrial fibrillation was defined. And in part, that is because there is some disagreement, particularly with cardiologists when they are brought onto the case, in terms of the length of time. You have these very, very brief runs. What does that really mean if you only have three beats? Do you commit someone to anticoagulation? And I wasn't sure whether or not that should

really be more defined in the definition or not.

But there wasn't a whole lot of generation of questions beyond that.

MEMBER R. BAUTISTA: Is the reliability data here the same as the prior study? It seems like it would be the same study as the last one we saw.

MEMBER WADDY: I'm sorry?

MEMBER R. BAUTISTA: The data we are using for reliability appears to be the same as the last study we actually analyzed. If you go to the old study, it is the same information they are giving here.

MEMBER WADDY: Yes, it is the same study.

MEMBER R. BAUTISTA: But one study actually looks at --

MS. YODICE: It is the same study, but it is still a different population that you are including.

MEMBER R. BAUTISTA: Right, but one

study actually looks at anticoagulation on discharge.

MEMBER WADDY: Right.

MEMBER R. BAUTISTA: But the other one is supposed to be on a few days after. So, it is supposed to be two different studies, then, right?

MS. YODICE: It is the same project, but not necessarily the same patients.

MEMBER R. BAUTISTA: Could you just make sure? Because the numbers seem pretty close to the last study. 0436 and 0241 are presenting with the same reliability data, I think.

They use four sites. They have 52 patient charts were eligible, which is the same as the last study. It seems pretty similar to the last study we looked at.

CO-CHAIR TIRSCHWELL: But these are different developers, are they not?

MEMBER WADDY: Yes, they are different developers.

CO-CHAIR TIRSCHWELL: So, it would be a different --

MEMBER R. BAUTISTA: But the question, though, being, what are we actually looking at now? I mean, one study here looks at anticoagulation on discharge, and the other one is anticoagulation --

MEMBER WADDY: It just says anticoagulation, but if you read within it, it says at hospital discharge.

MEMBER R. BAUTISTA: Well, the other one is on hospital day two.

MEMBER WADDY: That is the DVT, I think.

MEMBER R. BAUTISTA: No, no.

CO-CHAIR TIRSCHWELL: Let's make sure.

MEMBER KAPLITT: Both of the atrial fibrillation ones are at hospital discharge.

MEMBER WADDY: Just one includes the TIAs, and the other doesn't.

MEMBER R. BAUTISTA: It does not include TIAs.

MEMBER SUKO: I guess this is similar to

our other harmonizing measures between the AMA and the PCPI and The Joint Commission.

And just to clarify, this is based upon, it looks like, more physician Part B claims data than the facility data that The Joint Commission. Is that accurate? Okay.

MEMBER R. BAUTISTA: Go ahead. I am actually just looking at both studies right now.

CO-CHAIR KNOWLTON: Anything further?

(No response.)

We are ready to vote on reliability.

(Vote taken.)

MS. THEBERGE: We have 21.

Seven high, 15 moderate.

CO-CHAIR KNOWLTON: Validity.

MEMBER WADDY: And in terms of validity, there was agreement that at least it wasn't low, but it was split once again between high and moderate.

MEMBER BARRETT: I would just comment again that the exceptions are high on this

measure, 46 percent, and the definition of exceptions is an area I think for measure development to ensure that disparities in severe stroke that may be amenable to rehabilitation are not excluded inappropriately.

CO-CHAIR KNOWLTON: Okay. You raise that these are issues you want included as they consider these measures? You are not arguing for a particular position on this vote?

MEMBER BARRETT: Well, I am arguing that it may reduce the validity of the measure to some extent, but also stating for measure development for the future.

CO-CHAIR KNOWLTON: Okay. Thank you.

Yes, Michael?

MEMBER KAPLITT: Can I just clarify what that exception number is? Because you are talking about the reliability number? When you say it is the exception rate, that is what I am not understanding. Is this the --

CO-CHAIR KNOWLTON: Just a minute.

Ann, do you want to respond to Michael? Put your microphone on.

MEMBER BARRETT: Sure. As I understand, Salina may be able to give us some further guidance, but there are quite a few exceptions for either medical- or patient-based reasons.

But, you know, I could just see a patient who lives on his own and say, "Well, this patient is an exception because he doesn't have anyone else with him." But we don't know if that patient was actually incapable of taking the medication. They were confirmed to be reliable based on the documentation, I assume.

MEMBER WADDY: And actually, our group brought up that point, that there needs to be further clarity regarding what the physician reason may be for not including the patient. And that is one of the challenges. That is what I brought up in the other one, that patients who have a severe stroke, physicians

may not put them on anticoagulation -- or a large stroke, I should say -- may not put them on anticoagulation, and whether or not that really should be counted as a true exclusion or an exception.

There is a lot of argument regarding that, patients who go to rehabilitation and may fall, but have not had any falls. Is that really an acceptable patient -- you know, there is a question in terms of whether or not they can be given anticoagulation and how well that can be tested.

I mean, some physicians have conversations with their patients that are in-depth and some really don't.

CO-CHAIR KNOWLTON: Michael, does that answer your concern?

MEMBER KAPLITT: Well, I guess what I am wondering -- sort of, but if the reliability of the exceptions is high, which it is, doesn't that suggest that there's reasonable validity to the exceptions, even if the rate

is high?

I mean, we could all take exception; you know, we can all argue why certain exceptions could create problems. But let's say a theoretical group of practitioners, if 100 percent all agree that the exception is valid, then that is the best we can do, right? I mean, unless I am misunderstanding reliability.

CO-CHAIR TIRSCHWELL: The reliability just says that it was documented once and they can find it again when somebody else looks for the documentation.

MEMBER KAPLITT: Okay.

CO-CHAIR TIRSCHWELL: It doesn't say anything about the appropriateness.

MEMBER WADDY: I think it is more the validity question.

MEMBER KAPLITT: Okay. All right.

CO-CHAIR KNOWLTON: Hold it. We are not talking -- oh, I'm sorry, Michael. We are not talking reliability here; we are talking about

validity. And validity is whether the test or the measure is measuring what it purports to measure. Reliability would be whether it does it consistently.

So, we are talking about -- I assume you are saying that some of these exclusions would go to the validity question. Is that --

MEMBER KAPLITT: Correct.

CO-CHAIR KNOWLTON: Okay.

MEMBER WADDY: That is what I was hung up --

CO-CHAIR KNOWLTON: Yes, I think we are ping-ponging between two issues.

Anything else?

(No response.)

Okay. Let's vote on validity.

(Vote taken.)

MS. THEBERGE: Five high, 14 moderate, 3 low.

CO-CHAIR KNOWLTON: Okay. We will move on to usability. Salina?

MEMBER WADDY: So, in terms of

usability, it was, once again, split, but no one thought it was low. So, it was pretty much the same discussion: should there be more identification of absolute contraindications to anticoagulation versus an open-ended statement?

CO-CHAIR KNOWLTON: Any other thoughts?

(No response.)

Okay. Let's vote.

Oh, wait a minute. Jolynn, your hand is up.

Anything else?

(No response.)

Okay. Let's vote.

(Vote taken.)

MS. THEBERGE: Fourteen high, 8 moderate.

CO-CHAIR KNOWLTON: Okay.

MEMBER WADDY: And for feasibility, everyone agreed, for various reasons, that there were some challenges to the feasibility, but everyone rated it above low.

There was a question, once again, brought up of how patients were detected. And in the various types of settings are we missing a lot of patients, and then they are going home untreated?

CO-CHAIR KNOWLTON: Thoughts on feasibility? Risha?

MEMBER GIDWANI: Just so that I understand this properly, if a patient is discharged from a hospital and isn't prescribed the therapy, then goes to rehab and is prescribed the therapy in rehab, the physician who treated the patient on an inpatient basis would be dinged, but the rehab physician would not be dinged? Is that correct?

MEMBER WADDY: That is the way I understand it. There is no rehab because this is purely on discharge. Oh, I'm sorry. Yes, actually, it could be discharge from a rehabilitation facility.

MEMBER GIDWANI: No, no, according to

this, it is. The numerator says whether it is as an inpatient or a rehab facility.

MEMBER WADDY: Or a rehab facility.

MS. HANLEY: The measure is for discharge from the hospital. So, what you described is correct. The patient who is discharged from the hospital and not prescribed could fail the measure.

MEMBER WADDY: I'm sorry, who is talking? Oh, okay. Hi.

Except for in the numerator statement, it says discharged from a care setting, whether patient receives care in the ER or as an inpatient or rehabilitation facility.

So, I thought it included rehabilitation, in which case, yes, the first physician -- well, actually, if you went back, if the atrial fibrillation was not detected -- so, for example, it was in an ER setting; they only did a 12-lead EKG; they didn't detect it, sent the patient to rehab for whatever reason. And then, the rehab physician, say the patient

developed palpitations or whatnot, and then they detected it. Then, I am not sure how you would resolve --

MS. HANLEY: Yes, this measure, you are right, it is not as clear as it could be. And that is something that the group can work on. But I think the intent of this measure was for hospital, in acute-care hospital inpatient setting as opposed to rehab.

MS. JOHNSON: And just so you know where this is on the form, it is in Section 2a1.3435, Care Setting. It would tell you where they have specified this measure.

MEMBER KAPLITT: I mean, it does say specifically "acute-care setting". So, even in the rehab term, it would have to be acute rehab as part of the acute recovery. I mean, this wouldn't include somebody goes to rehab for two months or subacute, or anything like that, because that is not an acute care.

MS. HANLEY: Right.

MEMBER KAPLITT: I assume the intent

here is that a lot of places the patient recovers and then goes to the hospital's rehab center for another week or so as they are trying to be discharged, because there is a specific definition of acute rehab.

MEMBER BARRETT: And that is post-acute care, right?

MEMBER KAPLITT: Right.

MEMBER BARRETT: And so, the measure developers can confirm it is not used in the post-acute rehab setting, correct, in the post-acute care?

CO-CHAIR KNOWLTON: Okay. Dan?

MEMBER LABOVITZ: This is probably a non-issue, just a question. In terms of whose ICD-9 code identifies the a-fib, are we relying on the private practitioner? This is practitioner data. Or are we relying on the hospital which may not have spoken to the practitioner at all? Who is making the code?

MS. HANLEY: This is the physician-level billing.

MEMBER LABOVITZ: So, I worry that a lot of physicians say "stroke," and that is the only ICD-9 code they submit. And they are going to skate away scot-free.

Physicians, when they are submitting their private billing, aren't highly-motivated to put in a lot of ICD-9 codes. In fact, the disincentive is strong because most of it is paper billing, and they are just scrawling a number on a card.

I worry that, whereas a hospital wants a high level of acuity and will code everything, including warts -- so, I think that is a real validity problem.

MEMBER WADDY: We actually brought up that question within our Work Group, and what I was told was that this was still at the time of discharge and not in terms of when you go back and follow up with your primary physician.

MEMBER LABOVITZ: It is still the physician's ICD-9 code at the time of

discharge. You know, Dr. X, taking care of Patient Y, bills for stroke, discharges with a CPT code and an ICD-9 code. And maybe it is just one ICD-9 code.

MEMBER COONEY: It is a CPT II code, which are the specific codes for the performance improvement. So, I don't understand them well enough, but it is a whole different thing.

MEMBER LABOVITZ: I guess I am still confused.

MEMBER WADDY: The person who developed it, can they clarify?

DR. KATZAN: You know, I am speaking on behalf of the group, and they might actually have more input than I do. But I think this is based upon physician billing. So, it is correct. So, if somebody comes in and they just bill stroke, 434.01 or 436.01, and they don't bill for a-fib, then this patient would be in the denominator for this measure. I believe that that is an accurate statement.

MEMBER SUKO: Yes, I would echo Daniel's concerns.

And just a consideration for future testing: this is going to become an issue more and more on the physician and the hospital side, and there has got to be a way to test on the patient level what did the physician submit and what did the hospital submit in terms of testing and validity. And I would just throw that out to the developers. I recognize that that is not an easy thing to do, but we need to get to a point where we can figure that out.

Also, administratively, we are having the physician submit the exact same measures that the hospitals are submitting. So, we are creating a lot of work.

And so, as we develop this more and more, I mean, to the extent that we can streamline this, I think that is really important.

CO-CHAIR KNOWLTON: Other comments on

feasibility?

(No response.)

Okay. Let's vote.

Oh, I'm sorry, Mary.

MEMBER VAN DE KAMP: I was just going to say I think that it is difficult because so much of the outcome data currently is being driven off of ICD-9 codes. I mean, that is sort of the data-mining. I think that, unless there are other ways that we should recommend they look at this, unfortunately, the importance of ICD-9 coding and the hospital quality improvement has got to start around ICD-9 codes.

So, I think while it would be nice, I think we also have to recognize that much of the claims data is going to be driven off what is a code. And so, it may not be the developers' issue; it may be a practice issue.

CO-CHAIR KNOWLTON: Other comments?

(No response.)

Okay. Now let's vote.

(Vote taken.)

MS. THEBERGE: One high, 18 moderate, 2 low, 1 insufficient.

CO-CHAIR KNOWLTON: And the preliminary assessment, overall suitability for endorsement. Salina?

MEMBER WADDY: Everyone agreed to endorse it in our group, despite the challenges.

CO-CHAIR KNOWLTON: Any questions or comments?

(No response.)

Okay. Open it up.

(Vote taken.)

MS. THEBERGE: Twenty-one yes, 1 no.

CO-CHAIR KNOWLTON: I am going to suggest that we take a break, for all kinds of reasons. We will make it a 10-minute break. Does that work? Okay.

(Whereupon, the foregoing matter went off the record at 3:50 p.m. and resumed at 4:01 p.m.)

CO-CHAIR TIRSCHWELL: We are going to get started.

So, the next measure is screening for dysphagia. It is an AMA/PCPI measure, and Michael is going to get us started.

MEMBER KAPLITT: So that we all don't become numerators in the DVT prophylaxis outcome measure, I will try to move this along.

(Laughter.)

CO-CHAIR KNOWLTON: We are all using mechanical measures.

(Laughter.)

DR. SCHWAMM: You are assuming that you are getting treatment.

(Laughter.)

MEMBER KAPLITT: All right. So, this is re-endorsement of a measure that was originally endorsed in 2007 to do dysphagia screening in patients within, well, prior to PO intake following a stroke. And elsewhere, I think the ideal is within 24 hours of the

stroke, but the numerator statement is just patients for whom dysphagia screening was performed prior to PO intake, period.

In terms of the impact, I think that there was general agreement in the group that this was high impact. Clearly, they gave plenty of evidence to support the idea that aspiration from dysphagia is a major problem in stroke and represents a significant morbidity and even mortality, and probably the lion's share of pneumonias, for example, following stroke. And there were good papers to support it.

My only personal editorialization is that I found this to be an incredibly-convincing argument for a new measure. It would be nice -- and this was a common theme throughout a lot of these -- that if it is a measure that has been out for five years, you know, it would be nice in terms of impact to see like what data they have collected over the last five years among those that have used

this thing to see, is it making an impact?

You know, I think the potential impact is great, but it is still presented a little more theoretically than might be nice. But that is not really unique to this.

So, I think the impact, we all agreed, was high. I mean, it is important to do this.

CO-CHAIR TIRSCHWELL: Okay. Any comments about impact?

(No response.)

Let's go ahead and vote.

Wait. Wait. Now we can vote.

(Vote taken.)

One short. There we go.

MS. THEBERGE: Twenty high, 1 moderate.

MEMBER KAPLITT: Okay. Next, evidence?

CO-CHAIR TIRSCHWELL: Yes, go ahead to evidence.

MEMBER KAPLITT: So, there was complete agreement on this that the evidence was justifiable in terms of the subscores. There was variable agreement on high versus

moderate, but, overall, there was complete agreement.

And the basic issue, as with a lot of these things, is that it was felt that there were a lot of studies that were presented that justified this. It is just there weren't that many randomized controlled trials, but there were a lot of studies, and they were supportive. So, the evidence was felt to be high or acceptable, or whatever.

CO-CHAIR TIRSCHWELL: Great.

Any comments? Bill, was that a hand or that was just a yawn?

(Laughter.)

Okay. Let's go ahead and activate the voting then. Go ahead and vote.

(Vote taken.)

MS. THEBERGE: Twenty yes, 2 no.

CO-CHAIR TIRSCHWELL: Now to performance gap.

MEMBER KAPLITT: Right. So, in terms of performance gap, there was pretty good

evidence presented, one, from published literature, that dysphagia screening prior to feeding is all over the place, anywhere from 19 percent to 81 percent.

Again, this developer uses the CMS/PQRS reporting measures. They have for a lot of other things. And so, whatever we think of that, the performance gap based on that was 68.3 percent of patients did not meet the measure, but it seemed like the numbers were getting better in 2009, 2010. But, again, we all feel -- and I certainly feel -- that I think the PQRS overestimates, because these are people who are reporting because they get paid to report. So, I think that the performance gap, what they have documented is probably worse than that, and we all, I think, generally agreed on that in the group.

CO-CHAIR TIRSCHWELL: So, any other comments on performance gap?

(No response.)

Let's go ahead and activate the voting.

And now, go ahead and vote.

(Vote taken.)

MS. THEBERGE: Nineteen high, 3 moderate.

CO-CHAIR TIRSCHWELL: Okay. Moving on to scientific acceptability, first comes reliability.

MEMBER KAPLITT: Right. So, with reliability, just to make clear that the type of screening that they are referring to in all of this, which maybe I should have made clear earlier, can be a variety of things, including a video fluoroscopic swallow eval, fiber optic evaluation, a modified barium swallow, or a structured bedside evaluation. So, anywhere from extremely high-tech to more easily implemented.

And I think that the data that they presented was felt to show pretty high reliability or extremely high reliability with a good kappa score as well.

Again, this used, I think, the data

based on -- they used the four practice sites that I think we saw in some of the earlier things, where one was like a small group in the Midwest; one was a hospital in a city in the Southeast. I think it is that same group that we had sort of fleshed-out earlier. So, I think we already went through whatever issues there were with that.

And the only other thing is that the only exception, the only exclusion was death. Again, we could all argue about, well, if somebody died from aspiration, that is maybe not so great. But I don't think that that probably represents a significant enough number to be an issue.

And then, the exceptions were the usual exceptions, you know, documenting medical reason why you couldn't do the test for whatever reason. And they found, again, that there was 100 percent reliability in terms of the exceptions.

So, that is the reliability.

CO-CHAIR TIRSCHWELL: Yes, Salina?

MEMBER WADDY: So, I was just wondering -- I wasn't going to bring this up, but since you pointed out the various testing, why is this considered a screening for dysphagia rather than just dysphagia detection? Because once you go into fluoroscopic detection, then you are kind of beyond just a plain, old screening test.

MEMBER J. BAUTISTA: It is probably because it does include bedside screening as a valid measure or a valid action.

CO-CHAIR TIRSCHWELL: So, Jocelyn, activate your tent.

MEMBER J. BAUTISTA: So, even though they have various measures, they don't require the high level. So, just doing a bedside screening eval is enough to meet the measure, right.

MEMBER WADDY: And I am fine with the inclusion of the bedside screening test, but the actual title is that this is just a

screen. Is it just a screen? Or is it just any testing for dysphagia, is really my question, just really the overall title of it.

CO-CHAIR TIRSCHWELL: Does the developer want to comment?

DR. KATZAN: Yes. So, the minimum standard is a screen. And if they fail that screen, then they will go on to further testing to see if they are appropriate to take PO. But the minimum requirement is a screen.

But if people actually have -- and some hospitals do; some hospitals actually have the speech therapy department that can come to the bedside immediately. And in those cases, I mean, that would count as well.

CO-CHAIR TIRSCHWELL: So, it is anything at the screening-level-plus meets the criteria.

Any other? Yes, Mary?

MEMBER VAN DE KAMP: I think you bring a great point. I think it is just being in the field, I think that there is a big

difference between a screen and an evaluation. One is a physician engagement.

And so, I think that if it is just going to be a screen, they should leave it at that and not identify it as an evaluative process, because there is a very big difference in the field between a screen and an evaluation.

So, it is misleading, I think, to say that it is a screen. I think detection would be better, and you can describe it, the full range. Or you leave it at a screen and you allow for it to be a non-evaluative, because I do think this is a very confusing thing in the field already with the screening process and who can do a screen. Does it have to be a skilled clinician or can it be who does it all the time?

So, while it may seem small in light of this, when in actual practice, it is a really big discussion and gets into a lot of payment issues as well down the line for when a screen is to be done and when you need to have a

physician order.

CO-CHAIR TIRSCHWELL: I guess I don't understand how having a higher level really confuses it. It would be one thing if --

MEMBER VAN DE KAMP: Because I don't think the validity of a screen is the same level of the validity of a test. So, you could say I have a high validity with modified barium swallow or with a FEES. You might not have the same high validity with just a screen.

CO-CHAIR TIRSCHWELL: So, why don't we save that for the next point on validity?

MEMBER VAN DE KAMP: Okay. I will just make that point now, and I won't go back and say it again.

CO-CHAIR TIRSCHWELL: Yes.

Daniel?

MEMBER LABOVITZ: Just to add a comment, the business of keeping somebody NPO until you can get a FEES means sometimes you are not feeding a patient for three or four days. I

think that comes at a very, very high price. This may be one of the unintended consequences of having this sort of measure.

I think dysphagia screening, it has been a while, but, as I recall, it has been shown that centers that at least pay attention to this do better than centers that don't. But paying attention is much, much different than specifying sort of tests count and what don't, and that sort of thing.

CO-CHAIR TIRSCHWELL: Well, I think they go to great lengths to leave it pretty wide-open, probably reflecting the lack of evidence about what the best way to screen is.

I mean, at my center, to get this done, half the time you have to do it in the emergency department somehow because they are wanting to give them pills by mouth. So, it is a big deal; there is no question about it.

And I think, as a result, we have progressively lowered our definition for what becomes an adequate screen. And we might even

let the residents do it, which, you know, is a really questionable practice, I think.

(Laughter.)

But I think it is important. But I think it speaks more to the validity of this as a measure than reliability necessarily. I mean, the reliability, if you --

MEMBER WADDY: No, no, completely agree.

CO-CHAIR TIRSCHWELL: Okay. So, let's vote on reliability. Activate it. Go ahead and vote now.

(Vote taken.)

MS. THEBERGE: Fifteen high, 7 moderate.

CO-CHAIR TIRSCHWELL: Okay. It seems like we partially discussed the validity point already. Anything?

MEMBER KAPLITT: So, the validity was based, in part, on surveying of their expert panel, as well as, again, the CMS sort of PQRI reporting data.

All I would say to the point that was just discussed is, while I understand sort of

the point, so the consequence of it is let's say their measure only did bedside screening. Essentially, what you would be saying is that, if you had a medical reason why you felt that the patient needed a higher-end test, you would then be requiring that that bedside screen be done before that test is done, even though you know that that test needs to be done. Or else that patient would wind up getting excluded.

So, I mean, I don't personally see sort of the downside, unless they were to require that. I agree that the data -- so, in the validity testing, I don't see -- maybe the developer could add to this -- I don't see any -- it is not that granular where they break it down by type. But I don't see the downside.

MEMBER WADDY: Actually, I was more in favor. I mean, I think exactly what you said was what you related much more eloquently than I did, but I am more in favor of just changing the name of this item from "screening for

dysphagia" to something that actually reflects what is being done, right.

MEMBER KAPLITT: So, your point is that screening has a certain connotation to some people.

MEMBER WADDY: Yes.

MEMBER KAPLITT: But, technically, I mean, I guess the developer can deal with that, but, technically, as far as the document is concerned, they do define what they mean by screening in the numerator definition because the numerator says it is a screen, whatever, and then their definition of the numerator lists all these things. So, they have defined screening their way.

You're right, that might not be what the general public views it as, but --

CO-CHAIR TIRSCHWELL: And I mean, as far as validity goes, of the expert group, this one had more disagreement, more less-strong agreement than many of the other ones that the expert group did.

Ramon?

MEMBER R. BAUTISTA: I mean, I share the same sentiment. If they are going to get dinged for something like this, then he or she must take responsibility for screening the patient.

But, even among the group in this room, I am sure we are going to be screening these patients differently because we haven't been formally trained on how to screen dysphagia.

So, how do you resolve that? I mean, we all know it is important. But if you don't know how to do it, though, I am not sure what this measure really tells you then.

DR. KATZAN: Can I respond?

CO-CHAIR TIRSCHWELL: Yes, please, go ahead.

DR. KATZAN: So, your points are well-taken. And actually, the semantics that are used to define what is actually dysphagia screening versus formal bedside swallow, all that stuff is really a critical issue that is

confusing to many people. And so, changing the name is something that we really need to take into consideration, and I think we are all for.

Regarding the tools that are used, I mean, it is spelled out what tools there are. There is the 3-ounce water swallow test, the dysphagia screening test, and then you can do something that is more specific.

And one of the points that is made here is that a screening tool is approved by the hospital's speech pathology/language services. And so, there is some regimented system that the speech pathology/language group approves of that is used in a systematic fashion.

So, most hospitals, for instance, that are primary stroke centers, they have a tool that has been approved by that hospital that is done in a systematic, similar fashion, and actually documented as such prior to PO intake. And so, that is the thing; that is the point that we want to get across, that

tools could vary. And actually, a lot of people actually pass tools around in hospitals, share similar tools, but that the tool has to be or should be approved by that hospital.

CO-CHAIR TIRSCHWELL: There was a Get With The Guidelines measure. Is that NQF-approved currently?

Dr. Schwamm, do you know off the top of your head?

DR. SCHWAMM: So, there was a Joint Commission measure --

CO-CHAIR TIRSCHWELL: Sorry. Excuse me.

DR. SCHWAMM: -- that was not approved at the previous cycle, which is a hospital-based measure that complements this measure, I think largely because there was some confusion and concern about this issue around screening versus thorough evaluation.

What I would just say is that I think there is a bare minimum that all patients have some assessment prior to something being put

in their mouth. If you notice that the patient is so obviously at risk for swallowing impairment that you didn't perform the formal screen, but simply went ahead and ordered the speech consult, that we ought to give credit for that as opposed to excluding those patients from the denominator, which would be the alternative construct.

So, I actually support the current construct in terms of the idea that everyone deserves a screen. If you do a more thorough version of a screen, followed immediately, essentially, by a full evaluation, terrific. Not every hospital has the resources to do that, and not every patient needs that level of evaluation.

MEMBER WADDY: I am actually in favor of just calling it a dysphagia assessment and just leaving it at that.

CO-CHAIR TIRSCHWELL: Okay. Well, that will be a recommendation.

Any other comments on validity?

(No response.)

We are not allowed to, I think, just change it unilaterally.

(Laughter.)

She is in charge.

Any other comments on validity?

(No response.)

Let's go ahead and vote then.

(Vote taken.)

MS. THEBERGE: Six high, 14 moderate, 2 low.

CO-CHAIR TIRSCHWELL: Okay. So, moving on to usability.

MEMBER KAPLITT: I mean, there wasn't a whole lot to say about usability there. You know, the majority of the group felt that there was high usability of the measure itself, you know, that it is fairly easy to use and report on.

CO-CHAIR TIRSCHWELL: Any other comments?

MEMBER KAPLITT: It has been used for

five years.

CO-CHAIR TIRSCHWELL: Let's go ahead and activate, and now go ahead and vote.

(Vote taken.)

MS. THEBERGE: Seventeen high, 5 moderate.

CO-CHAIR TIRSCHWELL: Feasibility, Michael?

MEMBER KAPLITT: Right. Similarly, I mean, we can debate the feasibility of actually doing these various tests at various times. The feasibility of implementing the outcome measure, I think everybody felt largely was high and, you know, feasible.

CO-CHAIR TIRSCHWELL: Any other comments?

(No response.)

Let's go ahead and activate the voting, and go ahead and vote now.

(Vote taken.)

MS. THEBERGE: Fourteen high, 7 moderate, 1 low.

CO-CHAIR TIRSCHWELL: Okay. And then, overall suitability for endorsement. Any comments?

MEMBER J. BAUTISTA: Question?

CO-CHAIR TIRSCHWELL: Yes, Jocelyn.

MEMBER J. BAUTISTA: This doesn't sit well with me. So, if there is a similar measure that failed to get NQF endorsement, shouldn't that be a part of our discussion, why it failed?

MEMBER KAPLITT: Aren't we evaluating the thing before us? I mean, there could be a lot of reasons why it failed, right? If there is concern specifically about what we are seeing, then that is a different story.

MEMBER WADDY: They may have thought of something --

CO-CHAIR KNOWLTON: I was actually on the Committee on this issue. This was probably the most debated issue on that Committee because it rested upon who was going to do the screening and whether the screening

would be timely. Because, as it was originally configured, it had to be a specially speech and hearing person or physician.

And it went back and forth and back and forth with the developer. The language was finally compromised to allow the use of a screening tool.

And it was tied into the direct experience of a stroke survivor who got screened on Thanksgiving weekend, and there was nobody to undo the NPO hold until four days later. Okay? So, that became the poignant issue at the time of the debate.

But there was a hasty compromise that was crafted after a good hour debate on this issue. So, that is what happened, you know.

It is unusual for NQF to compromise like that, but it was because of the specific thing that occurred at that time.

CO-CHAIR TIRSCHWELL: I would also note that in the interval years -- I mean, that was

a number of years ago?

CO-CHAIR KNOWLTON: Yes.

CO-CHAIR TIRSCHWELL: There have been a number of reports that have done a little bit better job linking the dysphagia screening to pneumonia rates and things like that. So, the evidence base has probably improved a little bit and maybe adding to some more momentum at this meeting as opposed to a previous meeting.

CO-CHAIR KNOWLTON: There was no consideration that this wasn't valuable. It was, the way the measure was defined, it could only be done by two classes of people, and that got to be impractical on weekends and evenings, and so forth. That was the problem.

CO-CHAIR TIRSCHWELL: Dr. Schwamm?

DR. SCHWAMM: We would certainly welcome a recommendation from NQF that a hospital-based measure of this be resubmitted, because we do think it is important. So, we would be delighted to resubmit a companion, a similarly-formatted measure related to

hospital-level compliance.

CO-CHAIR TIRSCHWELL: Well, it seems like there is a lot of enthusiasm from the group. Although I don't know that we have the ability to officially endorse it, I think you have heard comments to support that for sure, and I would certainly support it.

MEMBER LABOVITZ: Just make sure you name it right.

DR. SCHWAMM: We will call it "The Do The Right Thing Measure".

(Laughter.)

CO-CHAIR TIRSCHWELL: Okay. So, overall suitability, can we vote on that now? Let's just go ahead and trigger that.

(Vote taken.)

MS. THEBERGE: Twenty-one yes, 1 no.

CO-CHAIR TIRSCHWELL: Okay. Excellent.

CO-CHAIR KNOWLTON: Okay. The next one is stroke education. Gwen?

MEMBER BUHR: Okay. So, this measure is measuring the people who have had ischemic or

hemorrhagic strokes and who themselves or their caregivers have been educated on five specific things: the activation of the emergency medical system, the need for followup after discharge, medications prescribed at discharge, risk factors for stroke, and warning signs and symptoms of stroke.

It is part of the bundle of measures that are used in The Joint Commission's hospital accreditation and disease-specific care certification programs.

So, on the impact, the data submitted do not specifically address education or stroke education particularly. It talks about how stroke is very common and disability from stroke. And it talks about that patient education programs in general are impactful. But it doesn't say anything about stroke education and, in particular, about stroke education on patient outcomes. So, that was our sort of issue with this measure.

CO-CHAIR KNOWLTON: I was actually going to comment on this. I was thinking if I am thinking in the right place.

Does anybody else have any comment on this? Risha?

MEMBER GIDWANI: Well, I was just reading. It says here -- I don't know if this, Gwen, speaks to your question about the lack of information -- but this is saying here that less than 10 percent of acute stroke admissions arrive at the ED within one hour of stroke symptom onset, and less than 25 percent arrive within three hours.

And the developers then linked that to the need for mass public education to increase awareness about the importance of seeking care as soon as possible.

CO-CHAIR KNOWLTON: But that is not the measure, is it?

MEMBER BUHR: Yes.

CO-CHAIR KNOWLTON: That is not the measure.

MEMBER BUHR: Right. So, I think that this is like one step away from even a process measure. They are saying that people aren't arriving in time for stroke. And so, then, they are assuming that they don't know, and that if they educate them, well, then they will arrive on time. So, it is like, you know, really many steps to get to where they want to go.

CO-CHAIR KNOWLTON: Do you have a comeback, Risha?

MEMBER GIDWANI: I agree, Gwen, with what you are saying. I just wonder if that might be a very difficult research study to operationalize and even to be able to measure that link. So, I wonder if we may need to make that leap of faith ourselves. I would consider it to be a reasonable leap of faith, but, of course, we all have to make that decision ourselves.

MEMBER WADDY: Actually, we have ongoing research at NIH regarding that. There is a

SWIFT study in order to try to get patients into the hospital in time.

Largely, there is a challenge with the stroke education programs in that they just measure whether or not you present someone with information, and whether or not they take in that information, not whether there is actually a behavior change that occurs because of it.

The second point I wanted to make is that this is really information at the level of information being provided to the patient. Unfortunately, a lot of patients who come in, they may either have aphasia or they may have some other cognitive --

CO-CHAIR KNOWLTON: It says caregiver.

MEMBER WADDY: It includes caregivers?

CO-CHAIR KNOWLTON: Yes, it does.

MEMBER WADDY: Oh, I'm sorry. Then, never mind. I will, then, stop there on that one. Thanks.

CO-CHAIR KNOWLTON: I am chairing this

one, but I do have a comment on it. And that is, in my experience, not only as a stroke patient, but also in what I have observed since, this ends up being a little plastic bag filled with a bunch of leaflets and documentation in the file, which is absolutely useless.

I was functional post-stroke, stroke patient. But this doesn't, in my view, capture whether they -- it just says they were discharged home with documentation that they were given this or their caregivers were given it, not that there was any teach-back or there was any "Do you understand?", any of those variables.

Bill, did you have a point?

MEMBER BARSAN: Yes, I would agree. And I agree with the others, too. I think there is a big gap between, gee, people don't know a lot about stroke and the fact that, if we educate people, we are going to change behavior, because I don't think the evidence

is there to do that. I am not sure I buy it.

CO-CHAIR KNOWLTON: Anybody else on the Committee? Yes, go ahead, Ann. And then, I will get to you, I promise.

MEMBER BARRETT: I would just comment that it is absolutely correct; the exclusions don't even include what language the patient speaks, whether or not they are illiterate. So, the materials can be given and the checkbox can be checked.

CO-CHAIR KNOWLTON: I don't know what language you could have used when I had my stroke. I was speaking in tongues.

(Laughter.)

But go ahead.

MEMBER BUHR: The other issue that we talked about is that it is for people who have had a stroke, and that the evidence that they give is for people who haven't had a stroke. The general population is not aware of this major symptoms of stroke. They don't present any evidence for what stroke survivors are

aware of or not. And the only people that this measure is requiring you to educate are people who have had a stroke.

CO-CHAIR TIRSCHWELL: And they are going home, right?

MEMBER BUHR: Yes.

CO-CHAIR KNOWLTON: Other members of the Committee? Bill? Dan? I'm sorry.

MEMBER LABOVITZ: The problem with education is you can do it badly or well. But even if you do it well, what we have learned with attempts at stroke education in patients who haven't had stroke, where you are educating the community, is that the learning fades extremely rapidly. So, it is gone in three to six months. Even if you can find a signal at three months, it is gone by six.

There has been a fair amount of effort done on this. That experience shapes my view on this measure because this measure doesn't present us with any data at all. So, I am just looking for other experiences.

I have no reason to think that the little packet, the little stroke packet that we give to our patients, or the little line that says, "Smoke cessation is a good idea if you smoke," which gets us 100 percent rating on our smoking score, does any good at all.

CO-CHAIR KNOWLTON: Salina?

MEMBER WADDY: Actually, I'm sorry. So, the comment that I should have been making previously was that, even if a patient has an aphasia or some other problem with retaining information or understanding information, they will still get a checkoff, according to this, even if they don't understand, for whatever reason, whether or not it is because it is not culturally-appropriate. It is just a checkoff.

CO-CHAIR KNOWLTON: That was Ann's point, yes.

Lee?

DR. SCHWAMM: Yes. So, I would say that I certainly recognize that the measure is not

the strongest measure in the set by any means. There is, as Dr. Waddy was implying, the SWIFT trial which was conducted by Bernadette Boden-Albala in Columbia, based on a cohort of multi-racial/ethnic populations.

The data is in press, but has not been published. So, I can only tell you what I know from speaking to the principal investigator.

All patients were randomized to receive either an educational intervention or usual care. And those who received the educational intervention were followed for the subsequent years. If they had a recurrent event in this captive community, they were, then, assessed as to whether or not they had increased their frequency of activating EMS coming in early and being eligible to receive t-PA. And the treatment group did show a statistically-significant improvement and a behavior change.

It is not published. It is not available to you. So, you can't include that

in your packet.

But I think all I would say is that, when this measure changed from simply provide education to actually specifying the five domains that needed to be addressed with the patient, there was huge pushback from hospitals because they found that actually onerous. So, not just a checkbox, but actually onerous, meaning they were doing something.

And my only concern would be that removing the measure sends a message that it is really maybe not important; whereas, strengthening the measure would be something that could actually lead to further change.

But I don't think the measure is causing any harm besides maybe doing education better than others, but I do think that it is critically-important. And when we don't require it, sites don't do it at all, nor do they document it, but they don't do it all.

So, I would just encourage you to think

about the fact that education in this domain is quite important, maybe not the best education being provided, documented, and established with teach-back, but it is still critically-important.

CO-CHAIR KNOWLTON: David?

CO-CHAIR TIRSCHWELL: I mean, I agree. Even if there is a little bit of a leap of faith here, I think that doing your best to providing education to stroke patients is a standard, just a tremendously-important standard of care.

My recollection of the SWIFT trial, though -- and I have only seen some peripheral things -- is that it was more than just a little packet at discharge, that their educational intervention was a bit more intensive.

DR. SCHWAMM: Yes.

CO-CHAIR TIRSCHWELL: And I am sure we are not doing that widely. If that supports a more intensive educational intervention, I

am all for it.

But I think I agree that this is a common-sense thing, and I agree that the leap to better outcomes for patients is really hard to demonstrate. But I also think that supporting the concept to providers that this is an important thing we want you to participate in, and hoping that hospitals, in good faith, try to do a great job of it, is something that is important.

CO-CHAIR KNOWLTON: Michael?

MEMBER KAPLITT: So, my concern about the harmful aspect of this is, when you say, "Well, it probably won't do any harm," is that if certain types of more intensive education can be helpful, and we have a measure that basically can be satisfied by, as you say, throwing a piece of paper in a plastic bag and saying on the way out the door, "Here, you're educated," the concern is that you will falsely conclude that education doesn't work because the vast majority are going to do

that. And the evidence of that is you are saying that, when there were attempts to mandate more intensive education, that there was pushback. There is pushback because it is hard.

DR. SCHWAMM: Just to clarify --

CO-CHAIR KNOWLTON: Well, don't reply yet.

DR. SCHWAMM: Oh, I'm sorry. Excuse me.

CO-CHAIR KNOWLTON: We will get back to you.

MEMBER KAPLITT: So, my concern is that, you know, unless there is data -- and I agree that there is a lot of common-sense aspect to all of this stuff, but unless there is data to support the measure as presented, which is that any type of education, you take a broad population of people, you know, hospitals, let them do whatever type of education they want, and show that that actually has some impact. And in the absence of a published paper that says specifically what was done to actually

achieve this difference, my concern is that it is not just, well, maybe it will help and maybe it will do nothing; it could be very harmful because of the false interpretation.

CO-CHAIR KNOWLTON: Okay, Lee, go ahead.

DR. SCHWAMM: So, I think I just want to make two points.

The first one is that the measure specifications as documented would not permit stuffing an envelope in the bag as meeting the measure specifications. So, the measure specifications actually address five specific domains and need to be individualized to the patient's specific conditions.

So, I think hospitals may not be following the guideline and may not be abstracting properly, and could be being deceitful as to what they provided, but, as specified, it is more rigorous than that.

CO-CHAIR KNOWLTON: Not what is before us, though. What is before us is that they or their caregivers, we are giving educational

material. It doesn't say spoken to. It doesn't say that material is culturally-specific or language-specific. It just says they were given educational material.

And I do believe, unless I misunderstand the measure -- and I would be happy to have you clarify that -- shoving the pamphlets in the plastic bag does that, would meet the measure. Am I wrong?

DR. SCHWAMM: Well, I think it depends on how you interpret those specifications. The specification says that it has to be relevant to the patient's specific condition. So, if you have atrial fibrillation, that needs to be discussed with you, and the medications that you are prescribed to treat that condition need to be linked to that as part of the education.

So, yes, I suppose you could have a pamphlet that addressed all five domains and said to the patient on the way out the door, "Your risks are covered in here. Read them."

So, I don't think, from my reading of it, that that would actually address the measure specifications, but we don't know how hospitals are abstracting that.

So, the answer is I can't reassure you that that is not happening. On the other hand, that is not how the measure is written. Just like I can't reassure you that hospitals are making up reasons for non-treatment with antithrombotic therapy or anticoagulation therapy, but we hope that hospitals are doing their best in good faith to comply with the measure.

I just want to say, I mean, I agree with you; it would be ideal to have an even stronger measure. I think part of the problem is we don't have data that tells us specifically which types of educational interventions have been proven to have an outcome link. So, a little bit, we lack the evidence to further refine the measure to make sure that it meets all of those elements.

But I just did want to address the original comment that this would be harmful.

CO-CHAIR KNOWLTON: Fred, you are all set?

Ramon?

MEMBER R. BAUTISTA: I mean, I agree with Michael's comment a while back. I mean, when we approve a measure, especially in the first get-go, the so-called leap-of-faith-type approval, I mean, after "X" amount of years, you need to actually come back to us and say, "Is it working or not?" I mean, to say that we don't have any new development at that point; that we have just got to keep going with what we have so far because we have nothing better, is I think, frankly, a bit of a copout.

In other words, it doesn't excuse not to get better. I mean, for all we know, we could do a couple of things better here with this measure, right, and just accept what we have from three years ago.

So, I mean, we think we probably should be doing more than just accepting something because there is no better data out there.

CO-CHAIR KNOWLTON: Therese?

MEMBER RICHMOND: You know, I am in Section 2a1.3, which is the measure specifications. I don't see where it says it is individualized, et cetera. I just see -- and maybe I am just not reading this correctly -- that there is documentation that they were given written instructions about this, about this, about this, about this. So, what I am reading is not what I think I am hearing you say.

MEMBER KAPLITT: We don't want to be unfair. You have got to tell us what pages what you said are on.

DR. SCHWAMM: Yes. So, we are looking to find that. Each of the five domains are specified in the Hospital Specification Manual, but give us a minute to try to identify that for you.

CO-CHAIR KNOWLTON: While you are looking, Risha?

MEMBER GIDWANI: I agree with actually all the points that have been made here. I think that these measures are far from perfect. I think it is right that they need to be addressed for cultural-sensitivity, literacy, and that data do need to be presented, that these measures have been in effect for a long time.

For me, though, the overarching concern is that, if we do not endorse this measure, the message that sends, in the absence of having any other education measures, is that education is not considered important from an NQF perspective. That, to me, is more dangerous.

I would rather see an endorsement with a requirement that The Joint Commission present data next time and make a more comprehensive and cogent measure than to see no measure at all.

MEMBER LABOVITZ: Just let me explain something, and Reva can kick me under the table or above the table.

We don't have the option of endorsement with a recommendation to do it in a particular way that forces it. Your only clout, if you want to change something, is by not approving the measure, which will bring about a change in the measure. Now we can debate whether that is good or that is bad, which may be to your point, but we don't have the option you suggested.

The problem, the dilemma you get into here is whether you let a measure that has been around for a long time -- everything you say here is being transcribed and it goes to everything is here. So, it doesn't go back, "Oh, they just rolled over and decided this was a bad idea. Education is a bad idea." It won't go that way. This entire debate will be heard or will be read. So, it doesn't go quite that way.

But there is still an element -- I am not denigrating your argument; I am just fleshing it out a little bit to say we don't act in vacuum. Everything you do here, when we said, "God Bless you" to Gwen when she sneezed, it is on the transcripts. So, it is a real-time thing.

Did you have something to add to that, Reva?

DR. WINKLER: The only thing, just in response, is really what we are asking you is to please use the criteria, apply the information that has been presented, and answer the question to your best ability of, does the measure information provided meet the criteria as we have described them? And you have had plenty of experience today knowing what those are. So, that is really the fundamental question we are asking of you.

CO-CHAIR KNOWLTON: Bill?

MEMBER BARSAN: So, let me take a little different tact. You know, I agree that

education is important. I don't disagree with that. I think that should be something that we think is important.

What I am thinking of is, okay, if I go back to my institution or another institution and talk to people who are responsible for carrying this out, and they say, "What's the evidence that this helps?", I mean, I would like to be able to say, "Yes, we have good evidence," not "Gee, we're just trying to get you to do the right thing, even though we don't have any evidence that this is effective." I would have a really hard time taking that stance.

I think it is really important that we stress things that are important, but that there is evidence to back them up and we don't just do them because we think it is a good thing.

CO-CHAIR KNOWLTON: Dan?

MEMBER LABOVITZ: I think we have been very focused on data where it exists and have

really grilled our developers and each other on what these data mean and how valid it is.

And now, here we are in a data-free zone. I think there is real harm in spending money, and it costs money to do proper education, it really does. It is real nursing time. It is real doctor time. It is time I spend in my office.

But mandating that money be spent where we have no known benefit, none, I think is really a failure to do proper triage, which is the situation we are in. We are all strapped for cash. I want to spend my cash where it really makes a difference. I have a lot of places to spend it.

And these measures become mandates. They are unfunded, but I have to do it and I have to spend a nurse to do it. I don't want to spend that money unless I know it is really paying off.

CO-CHAIR KNOWLTON: Salina?

MEMBER WADDY: I do think it is

important that physicians educate their patients as well as family members and caregivers. But how you do it is completely a new question. And considering this is a stroke education component, I think it has to live up to the same stringency and criteria that the other measures have.

I do think that -- and you can correct me if I am wrong from the NQF -- I mean, we can make a recommendation, right, regarding strengthening this -- what is this called? -- strengthening this measure.

DR. WINKLER: Are you trying to say you would like to see better measures of stroke education in the future?

MEMBER WADDY: Yes, in the near future.

DR. WINKLER: Fine. The near future.

MEMBER WADDY: But the way that it is written as-is, it is very limiting.

CO-CHAIR KNOWLTON: Greg?

MEMBER KAPINOS: Dr. Labovitz said we didn't have any data. The same way we --

CO-CHAIR KNOWLTON: Move closer to your microphone a little bit, Greg.

MEMBER KAPINOS: The same way earlier on we used some of the evidence that was not purely targeting stroke patients, for instance, for statins when there is only one trial, but the other trials were in a more broadened group of cardiovascular disease, couldn't we do the same with education. Like next time somebody submits a measure that is addressing education, if we broaden it to cardiovascular disease, I think the cardiology literature has some evidence that education has been very helpful in ameliorating outcome.

Or maybe there will be a little bit more data than the zero data that we have now. So, I am just saying, what is our threshold to extrapolate data from other slightly different diseases into fields when we don't have enough data targeting exactly the population that we are talking about?

MEMBER WADDY: May I answer?

CO-CHAIR KNOWLTON: Yes.

MEMBER WADDY: Actually, there are some ongoing stroke education behavior change interventions, one that was actually mentioned, that was mentioned in the writeup, "the Protect D.C.," that enrollment has now completed or -- yes, that enrollment now has been completed. And in the next year, that data will be out. The SWIFT study is out as well.

I am actually in charge of the initiative for the stroke prevention intervention research project in which there are multiple --

CO-CHAIR KNOWLTON: So, you are not going to let us drop it?

MEMBER WADDY: Huh?

CO-CHAIR KNOWLTON: You are not going to let us drop it?

MEMBER WADDY: Oh, no, absolutely not.

(Laughter.)

DR. WINKLER: Just to respond to Greg,

you have opened that door rather interestingly about looking to other areas. And NQF, not too long ago, removed the endorsement on the heart failure education measure, for exactly the same issues that have been raised here.

CO-CHAIR KNOWLTON: Michael?

MEMBER KAPLITT: I would also just add, for those concerned about message, two things. First, if we know that there is paper coming out that may have a more rigorous standard and that shows a positive result, and we wind up endorsing a less-rigorous standard with lesser data, I am worried about that message, No. 1.

And No. 2, I think that since the developer clearly has an ongoing interest in this, I don't think that this developer is going away and going to become disinterested in this subject.

If there were pushback from the hospitals, et cetera, for a more rigorous measure, I would argue that, if there is justifiable reason to withdraw endorsement,

that would arm the developer to do a better measure later on, rather than send the wrong message.

CO-CHAIR KNOWLTON: Everybody all in? Are we done?

(No response.)

Okay. We are voting on the impact of this particular measure as written. Open the voting.

(Vote taken.)

MS. THEBERGE: Two high, 4 moderate, 4 low, and 12 insufficient evidence.

CO-CHAIR KNOWLTON: We're done with that measure.

Okay. Next is me.

CO-CHAIR TIRSCHWELL: Okay. The next measure, Dave, NIH Stroke Scale Reported.

CO-CHAIR KNOWLTON: This is Measure 1955. It is the NIH Stroke Scale recorded measure, which measures the percentage of patients 18 or older with ischemic or non-specific stroke with an initial NIH Stroke

Scale recorded. It is a process measure endorsed by the American Heart Association and the American Stroke Association.

It is a graded neurological exam that assesses speech, language, recognition, and attention, visual field abnormalities, motor and sensory impairments, and ataxia. And it lists some exclusions that you can read.

The group voted to say that it was important to measure.

Do we have any comments or questions?

MEMBER KAPINOS: I am curious to know if anybody -- I mean, why didn't we bring up, also, the fact that there is at least higher reliability, I believe, with -- yes, I think it was higher reliability -- with the Modified NIH Stroke Scale?

So, one of the developer NIH folks here, maybe one of your friends, since the stroke communities are really small, can relook again at each of those items. And I think there are two or three items that were removed and one

that was modified for the sensory. Instead of 0-1-2, it is like 0 and 1.

So, Modified NIH Stroke Scale, according to the recent literature, is actually more reliable than the NIH Stroke Scale. So, maybe we should think about that before we endorse again --

CO-CHAIR TIRSCHWELL: Does the developer want to address that?

DR. SCHWAMM: Yes. So, Patrick Lyden did develop a modified version of the NIH Stroke Scale, largely in an effort to try to increase adoption and improve reliability. It has never really taken hold. And in fact, subsequent studies of the NIH Stroke Scale have shown very high rates of weighted kappa agreement.

And so, I think that there is good justification for continuing with the full NIH Stroke Scale, and all of the data we have about the use of the NIH Stroke Scale to predict mortality, functional outcomes, that

is all based on the full scale, not the modified scale.

CO-CHAIR TIRSCHWELL: Bill?

MEMBER BARSAN: I actually was part of the group that helped develop the original NIH Stroke Scale. I believe in that, and I think it is a great measure. And I do think it is very valuable in predicting outcome and doing all the other things.

I am just not sure -- I don't see the link with improved outcomes. It is good for predicting outcomes. It is certainly good for evaluating stroke patients. But where is the link there? And I am not sure where the evidence is on that.

CO-CHAIR KNOWLTON: Bill, we discussed that during our Working Group meeting, a phone call. And it was the use of this -- and I am not the clinician here -- but it was the use of this in making the determination regarding initiation of t-PA. So, the outcome -- am I saying all this right? -- so, the outcome was

tied to making that decision. It might not necessarily be the outcome in terms of the specific patient.

That was the argument. I am just reiterating the argument we made during our session.

MEMBER KAPINOS: But, actually, there is also data that -- I mean, there are a lot of neurologists that practice saying that you should not use the NIH Stroke Scale because, actually, you can have an NIH Stroke Scale of zero and still receive t-PA if you are a pianist and it is involving your right hand. And you can have an NIH Stroke Scale of 23-25 and many stroke neurologists will still give you t-PA, even if it listed as a warning, or whatever, in that long list. So, the NIH Stroke Scale by itself is not necessary to decide about t-PA.

CO-CHAIR TIRSCHWELL: I guess I would ask the developer if they want to respond to the impact that this measure would have.

DR. SCHWAMM: Yes. So, I think these are all very reasonable concerns. I think measuring the NIH Stroke Scale itself, in and of itself, doesn't change outcome, just like doing a CT scan, in and of itself, doesn't change outcome. But it enables decisionmaking that can have a profound impact on patient outcomes.

So, No. 1, whether you pick this or some other arbitrary scale -- and this is by far the one that is most commonly used -- it is necessary to assess stroke severity or stroke disability to determine eligibility for t-PA. Whether you have a strict cutoff or not, you need some assessment of the degree of disability.

I think the second thing is that we have good data to suggest that NIH Stroke Scales above a certain threshold, say 10, have a high predictive relationship to proximal arterial occlusion, and those patients benefit from transfer to centers that have stroke units and

more comprehensive stroke care.

And I think the last, and most important, issue, I think, is as it relates to risk adjustment for stroke outcomes. Without stroke severity measured on all patients, it is impossible to risk-adjust outcomes to understand if care is different at different centers. And no degree of propensity-matching or other approximate variables captures the power that the NIH Stroke Scale does in models both in Medicare patients, in the fee-for-service Medicare dataset, which is 65 and older, as well as models with patients ages 18 and above.

The NIH Stroke Scale by itself, with no other additional variables, has a C statistic of .83 for predicting mortality and outcome. So, it is a tremendously-powerful variable. I think we all know that intuitively.

All of the models that you will hear about, I think it is probably tomorrow, that look at risk adjustment for claims-based risk

adjustment, have no severity marker, and therefore, cannot adequately risk-adjust the way that we would all like them to.

So, having an NQF-endorsed measure related to stroke severity would permit CMS to collect NIH Stroke Scale scores and allow that to become part of an appropriate risk-adjustment report that could be used on all hospitals, not just those selective hospitals who take the effort to record an NIH Stroke Scale score.

CO-CHAIR TIRSCHWELL: I guess I will make one comment, and then Salina and Risha.

It seems like, if that is the purpose, then it should be included as a component of a mortality measure, and not necessarily standing on its own.

Salina, I think you were next.

MEMBER WADDY: Yes. So, my comments are largely along this same line.

First, if it is really for t-PA decisionmaking, we are already measuring t-PA.

And so, you can use whichever scale you feel is necessary in order to make that decision, and that is really the metric.

But I really did like your second point regarding being able to equalize across centers and determine whether or not there is increased severity. Whether or not CMS needs that information, and they should make that decision, but you certainly threw a wrinkle in there for me because of it. You rescued it.

(Laughter.)

DR. SCHWAMM: Am I not supposed to respond? Sorry.

MEMBER GIDWANI: I may drown that. I think the NIHSS is important. I can't speak to the clinical aspect of things, but I can speak to the risk-adjustment component.

While it is and can be an important factor for risk adjustment, I severely question the actual use of it for risk adjustment. If there is no billing code associated with the NIHSS score, then that

can't be used for the CMS models or for the vast majority of models that actually rely on ICD-9 billing codes.

Even if there was an ICD-9 billing code, the likelihood is that that billing code would just assess whether the NIHSS was administered or not, rather than provide the actual score associated with the NIHSS. And that latter component is what would be necessary for the risk adjustment.

So, if the purpose of this is really to use NIHSS for risk adjustment, I don't know that this measure actually sets the parameters up so that it can be utilized thusly.

CO-CHAIR TIRSCHWELL: Do you want to respond, Dr. Schwamm?

DR. SCHWAMM: Yes. So, I think this is the Catch-22. Everyone who builds risk models says, "Well, we can't do any better because we don't have a stroke severity marker. So, these models are good enough."

This is the first step to getting stroke

severity markers into the administrative claims database. I suppose this is appropriate to share with the Committee. CMS has approached us about the possibility of creating a code that would actually codify the NIH Stroke Scale value itself as the measure of adherence. So that the information related to severity would be coded in the administrative datasets, which would then allow appropriate risk adjustment.

So, we see this both as having immediate clinical benefit, which is assessing stroke severity is an important component just like assessing the need for rehabilitation services, just like assessing presence or absence of hemorrhage. It is a component that helps determine the level of care that is provided and raises concerns about risks or benefits of other therapies which we know are efficacious, like t-PA.

But it also sets the stage for the opportunity to really dramatically leverage

this incredibly-important variable which has, until now, been unavailable for large-scale assessments of performance in the United States.

CO-CHAIR TIRSCHWELL: Ramon?

MEMBER R. BAUTISTA: I may be paraphrasing you here. I really like the idea of standardizing the way strokes are maybe quantified. And I think of the mini mental status exam, for example, or the Klasco-Coleman Scale as having reached that level of wide use.

But aside from the fact that this is mainly for t-PA at this point, it is mainly a research, outside of the t-PA protocol, it is mainly a research endeavor. And my only fear is that it may be misused for things other than what it was really intended for.

CO-CHAIR TIRSCHWELL: Ann?

MEMBER BARRETT: I will just comment that, although I, too, am moved by the idea of a severity measure, it is very important, if

this is going to be used in some way across populations, and particularly for individuals, in predicting risk of disability or cost of care or determining these variables, that it be valid across different types of stroke. And, of course, it is known that the NIH Stroke Scale underrepresents right brain stroke, the severity of right brain stroke. And thus, those survivors who may have as much as twice as long clinical stay or other higher risk of falls may be systematically underrepresented, as well as those who have thrombotic stroke, and that may introduce health disparities in different populations.

CO-CHAIR TIRSCHWELL: Ramon again.

MEMBER R. BAUTISTA: And also, requiring non-neurologists, ER doctors, hospitalists, family practice docs, to perform an NIH Stroke Scale may be a problem in implementation down the road.

MEMBER BARSAN: Let me address that.

CO-CHAIR TIRSCHWELL: Bill, go ahead.

MEMBER BARSAN: Actually, when it was validated to begin with, it was tested on residents, attending physicians, emergency medicine docs, nurses, and it has good validity across all groups. So, there is really not an issue with one group being able to do it better than another.

MEMBER R. BAUTISTA: But this would be to require pretty much everybody who sees a stroke patient to do that.

CO-CHAIR TIRSCHWELL: Any other comments?

(No response.)

I think we should go ahead and open the voting on impact.

(Vote taken.)

MS. THEBERGE: Five high, 10 moderate, 3 low, 4 insufficient evidence.

CO-CHAIR TIRSCHWELL: Okay. Moving on to evidence.

CO-CHAIR KNOWLTON: Okay. As we look at evidence, the Committee -- only three of us

voted on this -- but all three found that there was evidence, although there was one rating of low.

We kind of discussed many of the issues that were raised in that, in the discussion by the Committee. So, I will only go further if you need it or if someone has a comment.

CO-CHAIR TIRSCHWELL: Any other comments about the amount of evidence supporting this?

(No response.)

Okay. Let's go ahead and open up, activate the voting. Don't vote. Now you can vote. If you voted early, you probably want to vote again.

(Vote taken.)

Still too short.

MS. THEBERGE: We are short two votes.

MEMBER KAPINOS: In hemorrhagic stroke, it would have dropped significantly the quality of the evidence for the measure if we included an NIH Stroke Scale for hemorrhagic stroke, too?

CO-CHAIR TIRSCHWELL: We are kind of post-vote here. It was pretty close and one missing. So, I think we probably need to do that again.

So, why don't you go ahead and make your comment? And Dr. Schwamm, we will let you respond.

MEMBER KAPINOS: I wanted to know if right now this measure is focusing only on the ischemic stroke patients to receive an NIH Stroke Scale. Why didn't we broaden that to also hemorrhagic stroke? Is the only reason the lack of evidence in hemorrhagic stroke?

DR. SCHWAMM: So, traditionally, it has not been a scale that has been applied in hemorrhagic stroke patients. But in Get With The Guidelines, we do collect it, and many sites report it for their hemorrhagic strokes.

And actually, we presented an abstract at last year's International Stroke Conference showing that it retains the same degree of predictive outcome capacity when you look at

a global model of stroke, so all stroke types together.

So, I think it is more a question of custom, and there is not a lot of data out there on its use within hemorrhagic populations, but it probably remains valid.

MEMBER KAPINOS: That's fine.

CO-CHAIR TIRSCHWELL: Any other comments?

(No response.)

I think we have to re-vote on the evidence. So, let's all try to wait for the -- okay. Now everybody raise your hands high.

(Re-vote taken.)

MS. THEBERGE: We are still at 20.

CO-CHAIR TIRSCHWELL: One more time, everybody.

MS. THEBERGE: You have to vote either 1 or 2.

There we go.

(Laughter.)

CO-CHAIR TIRSCHWELL: All right.

MS. THEBERGE: It is 10 yes and 12 no.

CO-CHAIR TIRSCHWELL: That was very close.

Okay. So, I guess we are done, then, discussing this measure at this point, and we are moving on to 0244. Sorry. We are going to change again. We are going to change to 0441 first, Stroke No. 10, "Assess for rehabilitation" from The Joint Commission.

And David?

MEMBER HACKNEY: All right. Let me first give a little disclaimer. I don't do stroke rehabilitation, and I was originally mistakenly assigned and then re-assigned to a more appropriate topic. So, I am still supposed to do this introduction, but I wasn't on the conference call. So, some of my comments may have been dealt with on the conference call, but I didn't get that from reading the summary. And the rest may come from the fact that, as I said, I don't do stroke rehabilitation.

So, anyway, the goal of this is to, I suppose the goal is to ensure that patients who are admitted for stroke are appropriately assessed for rehabilitation and to document that that has happened. So, this looks at patients who are admitted for stroke and the proportion of them who are either assessed for or receive rehabilitative services before they are discharged.

There are exclusion criteria to remove from consideration: patients are transferred to another hospital rather than discharged, left against medical advice, died, or were sent to hospice care. So, it attempts to get at that group of patients who, presumably, should be evaluated for rehabilitation.

So, going on to 1a, impact, stroke rehab is believed to be an important component of stroke care. It is not entirely clear to me that this measure really addresses that, for reasons we will get into as we move through.

But the vote for impact was strong

support for impact among the Work Group when they discussed it. So, maybe we should stop at that point.

CO-CHAIR TIRSCHWELL: Okay. Any comments on impact for assessment for stroke rehabilitation?

(No response.)

Okay. Let's go ahead and open the voting. And now you can go ahead and vote.

(Vote taken.)

MS. THEBERGE: Fifteen high, 7 moderate.

CO-CHAIR TIRSCHWELL: Okay. Moving on to evidence.

MEMBER HACKNEY: Okay. Next, I have performance gap.

CO-CHAIR TIRSCHWELL: Well, we do evidence first. We are doing them out of order. Sorry.

MEMBER HACKNEY: Oh, okay. All right.

CO-CHAIR TIRSCHWELL: We have been doing that all day.

(Laughter.)

MEMBER HACKNEY: Yes, I know, but my notes are in the order that they were when I started.

Evidence. So, evidence was rated high by everyone. My not being a stroke person, not being a rehab person, the evidence talked about the value of rehabilitation, but not about the impact of making the rehabilitation decision before discharge. And I didn't see that there was really any data on that specifically. You can understand the logic. Presumably, that is when the people are accessible and you can collect the data. But whether making that decision and evaluation in followup is acceptable or not, I have no opinion.

CO-CHAIR TIRSCHWELL: Any comments on the evidence?

(No response.)

Let's go ahead and vote on the evidence.

(Vote taken.)

MS. THEBERGE: Eighteen high -- 18 yes,

4 no.

CO-CHAIR TIRSCHWELL: Okay. And now, back to the performance-gap issue.

MEMBER HACKNEY: Okay. Performance gap, estimates of compliance are very high, in the high nineties, and they have been going up very slightly. They can't go up much; they were 97 percent in 2009 and up to 97.4 in 2011. So, there is not much room for performance improvement, but if this is important, then it may be important to make sure that people keep doing it.

I had some question about whether a compliance rate that high might be different than 100 percent only because of discordances in the coding of the charts rather than actual failure to do it. But there isn't room for much improvement by performance gap.

There is a discussion of disparities that says there are disparities by ethnic group in stroke outcome and to some extent in stroke treatment, but not about whether there

were disparities in assess for rehabilitation or getting rehabilitation services.

MEMBER WADDY: There are.

MEMBER HACKNEY: There are? Okay. That wasn't documented.

CO-CHAIR TIRSCHWELL: So, there are big disparities in rehab --

MEMBER WADDY: For African-Americans as well as Alaskan Natives, there are different reasons for that. Unfortunately, rehab as well is tied into insurance and how long your insurance for pay for it and whether or not they will pay for it at all, which it is a bit separate, but it is what it is.

CO-CHAIR TIRSCHWELL: Yes.

Any other comments on the presence of a performance gap and whether these very high ratings that The Joint Commission reports may not capture the whole picture, which is what I think what is being suggested?

(No response.)

Okay. Let's go ahead and activate the

voting, and go ahead and vote now.

(Vote taken.)

MS. THEBERGE: We are at 20. There we go.

Eleven high, 8 moderate, 3 low.

CO-CHAIR TIRSCHWELL: Okay. Great. So, then, we are on to 2a, which is reliability.

MEMBER HACKNEY: Okay. There was a feeling that the reliability of the measure was high. So, there wasn't a lot of debate about that.

CO-CHAIR TIRSCHWELL: Okay. Any comments on reliability?

(No response.)

Let's go ahead and activate the voting. Go ahead and vote now on reliability.

(Vote taken.)

MS. THEBERGE: Nineteen high, 3 moderate.

CO-CHAIR TIRSCHWELL: And validity, 2b.

MEMBER HACKNEY: There was also generally strong support with most people

voting high for validity. There was some concern, as is stressed throughout the document, that since these are voluntary reports, they might have a biased set of institutions that are reporting. So, the actual results might not be as good as suggested. But, basically, people supported the validity.

CO-CHAIR TIRSCHWELL: Comments on validity?

(No response.)

Let's go ahead and activate the voting. Go ahead and vote now.

(Vote taken.)

MS. THEBERGE: Seventeen high, 5 moderate.

CO-CHAIR TIRSCHWELL: Usability.

MEMBER HACKNEY: There was strong support about usability. It has been widely used, and it is believed to be easy to interpret. So, I don't think there were any concerns raised about usability.

CO-CHAIR TIRSCHWELL: Any comments?

(No response.)

Let's go ahead and activate. Now go ahead and vote on usability.

(Vote taken.)

MS. THEBERGE: Twenty-one high, 1 moderate.

CO-CHAIR TIRSCHWELL: Feasibility.

MEMBER HACKNEY: There were some concerns about feasibility because this requires abstracting data from the medical record, and not all that information is necessarily in the electronic health record. And there were some concerns about how reliably you would get it all out.

On the other hand, the report indicates that they have data from an enormous number of hospitals. So, it would certainly suggest that it is practical for -- they didn't address what proportion of stroke patients are represented, but it is such a large number of places reporting it, that I would suggest that

that alone is proof that it is feasible.

CO-CHAIR TIRSCHWELL: Any further comments on feasibility?

(No response.)

Activate the voting, and vote on feasibility.

(Vote taken.)

MS. THEBERGE: Twelve high, 9 moderate, 1 low.

CO-CHAIR TIRSCHWELL: And then, finally, overall suitability for endorsement. Any further comments, David?

MEMBER HACKNEY: Everyone endorsed it.

CO-CHAIR TIRSCHWELL: Let's go ahead and open the voting.

(Vote taken.)

MS. THEBERGE: Twenty-one yes, 1 no.

CO-CHAIR TIRSCHWELL: All right. I guess we should go back now to 0244.

Do you want to do this one, Dave?

Also consideration of rehab services. Also, Dr. David Hackney.

MEMBER HACKNEY: So, this is the same preface as the last one. I won't repeat it.

Here, unfortunately, there were only two people on the Work Group call. So, we only have two sets of votes, but I will proceed.

This is a very similar measure to the last one, slightly different definitions of who included. And this is looking at patients in whom it is ordered, rehabilitation is either ordered or documented that such services are not indicated, as opposed to the prior one which was assessed or received. There are no exclusions.

There was a general feeling that this was important, at least that stroke rehabilitation was important. The same questions about whether documenting that this was ordered as an inpatient is critical, they didn't have evidence, but they believed that rehabilitation itself was important.

CO-CHAIR KNOWLTON: Any discussion on impact?

(No response.)

Okay. Let's vote.

MEMBER KAPLITT: Can I just ask a quick question?

CO-CHAIR KNOWLTON: Okay. I'm sorry.

MEMBER KAPLITT: I am sorry, I know it is late.

But, as written, I mean, is it possible that someone could order, let's say, speech therapy for someone who has no speech problem, but has, let's say, a paralyzed arm, and that would count? Am I misreading this?

MEMBER HACKNEY: I believe so if they wanted to do that.

MEMBER KAPLITT: Because my numerator, I mean, the numerator that I am reading --

MEMBER HACKNEY: Do you mean does it say "appropriate"?

MEMBER KAPLITT: Because as opposed to the last one, where it was that they were assessed and rehab was sort of based on the assessment, presumably, here it is based on

the order, as you said, right?

MEMBER HACKNEY: Yes. You mean if somebody could just default order that on every single patient, no matter what the nature of their problem was.

MEMBER KAPLITT: Something like that. Or they make a mistake and it still counts. Or they whatever. I mean, I am just -- am I missing this?

CO-CHAIR KNOWLTON: Well, at some point, you have some clinical judgment. And, yes, could clinical judgment be abused --

MEMBER KAPLITT: Well, no, but I disagree because the last one was more relevant to what is going on with the patient. They are assessed for whatever their problem is. Here it is just based on an order. There is no indication of relevance of that order to the patient.

MEMBER HACKNEY: The presumption was that the assessment preceded the order, but, I mean, you know, technically, I guess you

could order it without assessing.

MEMBER KAPLITT: I mean, I am just looking at this for the first time.

MEMBER HACKNEY: Yes.

MEMBER KAPLITT: I am just asking. If I am wrong about this, you know, then tell me I am wrong.

MEMBER KAPINOS: I agree there are institutions that have order sets for every stroke patient regardless of ischemic, hemorrhagic, or whatever. And so, at my institution every stroke patient gets the whole PT/speech therapy order. It is a blanket order for everybody. So, yes, patients with no speech deficits, I mean the speech therapist will receive the order in the computer. And when they come at the bedside, then they will ask, "Do you really want me to see that patient?" And then, we realize, "Oh, no, actually, cancel that order."

CO-CHAIR KNOWLTON: Anybody else?

MEMBER COONEY: I was just going to

comment that I think the reason that it is based on the order is because it is a clinician-specific measure, and that is what the physician does, is they order it. They don't actually perform it. So, that would be my presumption.

MS. HANLEY: Yes, I was just going to say that the intent is that the physician determine what is appropriate for that patient, and then take that next step to order it, if indicated, or document that it is not indicated.

MEMBER KAPLITT: I have no problem going on faith. I am just saying that that is not what it says. You could do an order where it says that it is disability-appropriate or based on the physician evaluation; it gets ordered. I mean, there are ways to do it that are physician-specific that still makes it clear that they have had to actually make the order relevant to the patient, because you are giving a broad array of things they could

order that are not all equivalent.

But that's fine. I am happy to take it on faith. If that is what everybody else thinks, that's fine.

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

MEMBER SUKO: Well, I just wanted to respond because at one point in my organization I wore the compliance hat. And so, typically, a pre-checked order for speech, OT, and PT, at least the way we interpreted that was that that was a potential compliance risk.

So, CMS would -- I mean, it gets to protecting, to creating that actual assessment, and that there is actually some physician judgment in the services that are being ordered. Now I am not going to say that every organization is going to interpret it as conservatively as we did, but that is from the compliance world what they would probably say about it.

CO-CHAIR KNOWLTON: So, change your order sets, Greg.

Anything further?

(No response.)

Let's vote on the impact.

(Vote taken.)

MS. THEBERGE: Thirteen high, 9 moderate.

CO-CHAIR KNOWLTON: Moving on to evidence, David.

MEMBER HACKNEY: So, there was support for the idea that there is evidence, an evidence base to this.

Do you want to do the subgroups of evidence or just that statement first?

CO-CHAIR KNOWLTON: Just the evidence, yes or no.

MEMBER HACKNEY: Yes or no, the feeling was, yes, there is evidence on the subject.

CO-CHAIR KNOWLTON: Comments? Thoughts?

(No response.)

Let's vote on evidence, up or down.

(Vote taken.)

MS. THEBERGE: That is 20 yes, 2 no.

CO-CHAIR KNOWLTON: Okay. Now we move on to performance gap.

MEMBER HACKNEY: Okay. There was mixed feeling about performance gap, and I think the reason for that is they weren't happy with the relationship between the performance-gap evidence about how many stroke survivors get stroke rehabilitation and this particular measure, because there was good evidence that a lot of stroke survivors don't get stroke rehabilitation.

But the feeling, I gather, not having been on the call, was that it wasn't clear how important the stroke rehab was to the people who weren't getting it, and it wasn't clear whether the problem was that they didn't get assessed or ordered while they were inpatient. So, it wasn't exactly on point to this measure, but there is certainly evidence that a lot of people have strokes and don't get

rehab.

CO-CHAIR KNOWLTON: Comments on the gap?

(No response.)

Okay. Let's vote.

(Vote taken.)

MS. THEBERGE: Eleven high, 9 moderate, 1 low, 1 insufficient.

CO-CHAIR KNOWLTON: Okay. Reliability.

MEMBER HACKNEY: The two people thought that the reliability was high. When they did repeat abstraction from the medical charts, they got very good reproducibility.

CO-CHAIR KNOWLTON: Comments on reliability?

(No response.)

Let's vote.

(Vote taken.)

MS. THEBERGE: Sixteen high, 6 moderate.

CO-CHAIR KNOWLTON: Validity.

MEMBER HACKNEY: That was assessed by expert consensus evaluating face validity, and there was very high agreement in that it was

a valid measure.

CO-CHAIR KNOWLTON: Comments?

(No response.)

Okay. Let's vote.

(Vote taken.)

MS. THEBERGE: Thirteen high, 9 moderate.

CO-CHAIR KNOWLTON: Usability.

MEMBER HACKNEY: This is already in use in several large QI programs that we have been discussing here. So, there is a feeling that it has demonstrated usability.

CO-CHAIR KNOWLTON: Salina?

MEMBER WADDY: Was there any discussion regarding -- and I guess this actually gets more to both usability as well as feasibility -- but some of these small hospitals, I mean, they take care of a lot of patients around the country. And for them to have PT, OT, speech, or even rehab services, I mean, a lot of them just don't have it.

And I am not saying to get rid of this

measure because of it, but there is a real reality when this measure is for whether or not the service was ordered at or prior to discharge. So, they may only have outpatient rehabilitation that may or may not happen.

MEMBER HACKNEY: This measure, somebody would be in compliance if they ordered, while the patient was an inpatient they ordered outpatient rehab.

MEMBER WADDY: Okay.

MEMBER HACKNEY: You would be in compliance with that, whether it actually happened or not.

MEMBER WADDY: Right.

MEMBER HACKNEY: If that addresses your question --

MEMBER WADDY: Yes, I mean, I can certainly think of -- I mean, I grew up in Alabama. When you are in the middle of western Alabama, you can order outpatient rehab all you want, but -- (laughter) -- Alabama, Mississippi, it is a huge gap.

MEMBER HACKNEY: Yes.

MEMBER WADDY: I mean, unless you have somebody that can actually take you 300 miles away, it is not going to happen.

MEMBER HACKNEY: Yes, I think that is true. I don't know if there is anything in terms of this is a performance measure --

MEMBER WADDY: The middle of Alaska --

MEMBER HACKNEY: Yes.

MEMBER KAPLITT: Correct me if I am wrong, but just usability I thought relates to the usability of the measuring instrument, right, not to the intervention? Right? It is supposed to relate to --

MEMBER HACKNEY: Right.

MEMBER KAPLITT: -- how usability the measuring instrument is here. If people can't do the actual intervention because of -- you know, that is part of the deal, is that it will, hopefully, help get more resources, or whatever.

CO-CHAIR TIRSCHWELL: Perhaps, Salina,

your point is more to validity and whether there should be some exclusion for I work in the boonies and it is not possible.

(Laughter.)

MEMBER WADDY: I think that it is so important that it is not okay to just exclude it, but whether or not things that are now coming down the line are using more technology such as home telerehab or things like that, that really aren't built into the system, I mean, I have been in the middle of Native land or the frontier of Alaska, and you have got to travel by dogsled in the middle of the winter or get on one of those little, bitty planes that are death traps.

Oh, I should probably strike that.

(Laughter.)

But, no, those little, bitty planes, I won't get on them.

But, I mean, there is a reality to it, that they are not going to put their family member who has an NIH Stroke Scale score of 16

and send them that far away, where no one can visit and no one can see them, and they don't know what is going on at all.

MEMBER HACKNEY: If I understand it, this is more a measure of the behavior of the physicians who are managing the patient while they are an inpatient. And that, the situation you described, would be out of their hands. I can see the concern that a hospital could show extremely high compliance even though very few of their patients are actually getting any rehab.

MEMBER WADDY: Exactly. Right. It won't mean anything.

MEMBER HACKNEY: But that would require a totally different measure that is looking at rehab achieved as opposed to -- which would be done by people other than those who are managing them during their inpatient admission.

So, I agree it is a consideration.

CO-CHAIR KNOWLTON: We can just add

Alabama and Mississippi to the exclusion. Right?

(Laughter.)

MEMBER KAPINOS: Doesn't it go along with what we have been discussing like for all those measures? It sounds like we have been bashing on like the fact that, if a measure is just to document something or to put an order, it won't have any impact on the outcome directly. Therefore, we should not vote for those measures.

And now, like we are going along with -- yes. So, I think this measure is the same as the other ones. If you are only documenting that you ordered something or just checking a checkbox, I don't think that is a good measure. Right? Or am I wrong?

CO-CHAIR KNOWLTON: Well, I don't know if you are right or wrong. But, in terms of how it is used, very often, this becomes documentation of the need in an area that, without this measure, you wouldn't have.

So, while I agree with what Salina said, it is a problem in certain areas, and there are certainly disparity issues that surround it, you never get at those disparity issues if you don't have some measure. And you don't have the forcing function of the doc saying, "This is what should happen."

See, the other problem you get into is, then, you don't write them in Alabama because you know it is 300 miles away, and that is a problem, too.

Go ahead. Go ahead.

MEMBER WADDY: I completely agree with that, but that is why I --

CO-CHAIR KNOWLTON: That is why I said, "Go ahead." I knew you were going to agree.

MEMBER WADDY: Huh?

CO-CHAIR KNOWLTON: That is why I said you could go ahead. I knew you were going to agree.

(Laughter.)

MEMBER WADDY: I am actually proposing

or suggesting or thinking about not getting rid of this measure, but actually having an additional measure that captures whether or not patients actually receive those services.

The same, it is not good enough just to order a medication if it never makes it into the patient's body because t-PA was not at that hospital. But, yes, we went ahead and ordered it; that is awesome.

CO-CHAIR KNOWLTON: Well, you have just made that recommendation because it was on the record.

Mary?

MEMBER VAN DE KAMP: I was just going to say that I think that in this case, while there are exceptions to where you can access therapy -- and we certainly see that -- it is more for the overall impact of ordering the rehabilitation services.

But, to your point, I think when you get out of the hospital, it is difficult. And I think everyone is having that with the post-

acute arena: how do you really follow the patient once they have left?

But I think, while those are the exception, the majority still are impacted by the ordering appropriateness.

CO-CHAIR KNOWLTON: Greg, do you have a point?

MEMBER KAPINOS: Didn't we just endorse actually a measure that was actually slightly better, and like getting rid of all this discussion that we just --

CO-CHAIR KNOWLTON: Yes, but that is for later. We are going to reconcile. We are going to harmonize.

You look at the measure standing alone. Then, you look at it compared. We did a number of measures that measure the same thing.

Anything else on this? This is on usability.

(No response.)

Okay. Let's vote.

(Vote taken.)

MS. THEBERGE: Sixteen high, 6 moderate.

CO-CHAIR KNOWLTON: Feasibility.

MEMBER HACKNEY: The two Work Group members on the call thought it was feasible, and the primary evidence is that it is very widely-used. There was some concern raised that not all the information is necessarily easily extractable from the electronic health record, but because places have been using it successfully, it looks like the measure itself is feasible.

CO-CHAIR KNOWLTON: Comments?

(No response.)

Okay. Let's vote.

(Vote taken.)

MS. THEBERGE: Seventeen high, 4 moderate, 1 low.

CO-CHAIR KNOWLTON: And finally, the overall suitability. David?

MEMBER HACKNEY: I'm sorry. The Work Group people found it suitable for

endorsement.

CO-CHAIR KNOWLTON: Comments?

(No response.)

Okay. Let's vote.

You were quick on the flip there, Suzanne.

(Vote taken.)

MS. THEBERGE: We are at 20.

Twenty-one yes, 1 no.

CO-CHAIR TIRSCHWELL: Okay. That is the end of the measures, although there is still the consideration of candidate measures.

Would people like to take a five-minute bio break or just power through and get out of town here? Bio break? One person, two people. That is a minority. Why don't you all just excuse yourselves for a minute and we will power through? Oh, more than two.

(Laughter.)

All right, five minutes. Come back in five minutes.

(Whereupon, the foregoing matter went

off the record at 5:36 p.m. and resumed at 5:42 p.m.)

MS. JOHNSON: I know everybody is going to be very disappointed to find out that we are not going to have two hours to talk about related and competing this evening.

(Laughter.)

So, what I am going to do is give you a one-minute spiel, I think, on what related and competing is. And I think you guys already know because it has come up many, many times as you have talked about different measures and such.

And while I am doing that, Jessica is going to be walking around giving you a handout. And you can use it as homework tonight, if you choose to do so, or we will look at it tomorrow, so however you want to do that.

So, I don't remember if this is on your Quick Guide or not, not exactly. But, basically -- yes, more stuff -- basically, we

want to think about different measures, measures that are in the project that you are in, so in the neurology project. And we also bring in potentially other measures that you have not looked at but may be in our portfolio that have been endorsed previously. Basically, we think about them at a conceptual level.

So, what we do is we pretty much think about the patient population, so the denominator, and we think about the measure focus. What is being measured? That is usually the numerator.

So, what you want to think about is, conceptually, do two or more measures look at the same measure focus? So, are they measuring the same thing? And/or are they measuring the same population? Okay?

So, let's do, hypothetically, if we have two measures and one of them is looking at -- I will make this up -- mortality amongst patients with stroke 18 and older. And

another looks at mortality among stroke patients 65 and older.

So, in that example, both of them are looking at stroke mortality. So, that would be the measure focus, and that would be the same, right? Both of them would be looking at mortality.

But the populations are a little different. But, really, the population, conceptually, are stroke patients. The fact that one set is 18 and older and the other set is 65 and older is a detail, but, conceptually, you are still looking at stroke patients.

So, you have to kind of get yourself a little bit out of being very concrete about the specs and work backwards and think about conceptual. Are they the same or are they different?

Well, we do that work for you. So, we present to you the ones that we think are related or competing. Sometimes they are

almost both. And sometimes, conceptually, you can argue that one is one or the other.

But, at the end of the day, what we ask you to do is, if measures are competing, we want you to try, if you can, to select the best measure. Because if they are competing, basically, what you are saying is I am measuring the same basic thing in the same basic population. And why are there two measures when there could be one? So, if possible, we would ask you to pick the superior measure. If you can't do that, then we would ask you to justify the reasons why you think there is a need for two measures.

Now in orientation we went very quickly through that. I will go through that a little bit more in detail tomorrow. But if you happen to want to look over the orientation slides, there is a little flowchart that gives you the steps to think about what you do. In other words, how do you pick a superior measure?

For related measures, either your target population is different or your measure focus is different, but they are related. So, when you have related measures, what you would like to have are measures that are as similar in their specifications as they can be.

So, that is where we start talking about harmonization. So, to go back to my example, if you had stroke mortality measure, two of them, it is not a great example, but you would ask yourself, is there any way that we could do something so you didn't have two measures and the only difference is age? Okay?

So, to give you a flavor of tomorrow, we have already heard that some of the measures have TIA in the denominator and some don't. That would be a place where you would start thinking about, is there a way to harmonize those measures, so that they are very similar in their specs? And again, if yes -- or you can justify the differences.

It is probably one of the harder things

that we ask you to do, thinking about related/competing. So, that is why (a) we are not going to do it, because we don't have time, and (b) I think if you are like me, you might not have the brain space to be able to do that this afternoon.

But go ahead and ask questions.

MEMBER KAPLITT: That's fine. If you can clarify one thing? We talked today, we had several measures today that were similar, but one was measuring the physician -- it's on (referring to the microphone).

MS. JOHNSON: Oh, sorry. That's fine.

MEMBER KAPLITT: We had several measures today where they were measuring basically the same thing, but one was measuring the physician, let's say the order, and one was measuring the hospital doing it. So, what category would that fall under? If they are measuring the same thing in the end, or they are theoretically measuring the same thing, but they are measuring whether the physician

is doing it versus whether the hospital is accomplishing it, how is that dealt with?

MS. JOHNSON: We would call that, according to NQF definitions, we would call those competing measures. And if possible, we would ask you to state if you think one is superior over the other.

Now, often, in those particular cases, having a different level of analysis is often a reason that Steering Committees would say that is a justifiable reason to have two measures, because one looks at clinician and one looks at facility. So, you would have to decide for yourself whether you think that is enough of a difference to justify having two.

CO-CHAIR TIRSCHWELL: Can we still suggest harmonization if they are still -- like we need the clinician versus the hospital, but the criteria are a little different, and we would really like the criteria to be the same?

MS. JOHNSON: Yes, you can make

recommendations. They are like every other recommendation that you have made today.

CO-CHAIR TIRSCHWELL: Non-binding.

(Laughter.)

MS. JOHNSON: We will put that in the report. There is no mechanism to assure that your recommendation will be taken. That said, developers, of course, do try to accommodate, when they can, but a lot of changes may make a lot of difference in terms of whether their testing would still apply. So, it can be a very big question and a big deal, even something that seems small.

So, really, your only stick, if you will, is to say, okay, if there is no opportunity at all for harmonization, then we could actually reverse our recommendation for endorsement. So, what you have done today is you have recommended a lot of measures as suitable for endorsement, but that is kind of the pre-related/competing. So, things that you have recommended today could still go down

tomorrow if you think that, for example, one is less superior than another.

CO-CHAIR TIRSCHWELL: Gail?

MEMBER COONEY: Most of the ones that are similar to what Michael was describing are in the physician/the hospital setting. And if you just look at the ones on the front of the page, it is prescribing antithrombotic therapy at discharge, which is something only a physician can do. But the way the measure was constructed and The Joint Commission materials, it was more consistent and supportive, and yet, it is really a physician -- I mean, can the hospital affect whether or not a physician writes the prescription? I mean, does that come into play at all when you are trying to sort these out?

MS. JOHNSON: I think it would come into play, yes, although, hopefully, you thought about some of these issues when you went through the measure the first time. So, you know, what you just asked is, I believe, a

validity question. Is this measure really accurate and can that really distinguish quality between facilities? So far, if you had that, then you have, by your vote, said that you believe that it is a valid measure for facilities.

MEMBER LABOVITZ: Speaking as a physician, I can say hospitals make us do all kinds of stuff.

(Laughter.)

MEMBER WADDY: I was going to say the same thing, too, but it really is, overall, what is going on in that hospital. I said previously there are three different practices in the Fairfax INOVA Hospital. And so, if you really wanted to know is everyone being treated that comes into that hospital with antithrombotics, that is one question. But if you went to individual practitioners and individual practices, they may have or may not have different practices.

MS. JOHNSON: So, if you take a peek at

the homework sheets that we sent around, there are two overarching questions just to kind of keep in the back of your mind as you think about related/competing tomorrow. One is, do we need all of these measures? And then, the second, if so, have they been harmonized as much as possible? So, that is what we are trying to get to with the related and competing conversation.

I think if we have no more questions and we are almost out of time, we still have to open this up to public comment.

You're on the ball. Does anybody have any more questions before we end and hand it over?

CO-CHAIR KNOWLTON: We have to open for NQF comment.

I also want to highlight we have an early start tomorrow, which will be especially important to us getting out tomorrow afternoon. Having sat through a number of these, I can warn you that really is a risk.

So, please be timely tomorrow.

Can we have them open the phone line for comment, if there is any public comment?

MS. JOHNSON: Amy, can you open the lines for public comment?

OPERATOR: Okay. If you would like to leave a comment or have a question, please press *1.

(No response.)

There are no comments at this time.

CO-CHAIR KNOWLTON: Anybody in the public seated here have any comments?

(No response.)

Seeing none, we are adjourned.

(Whereupon, at 5:54 p.m., the meeting was adjourned for the day, to reconvene the following day, Thursday, June 21, 2011.)

G

N

X