NATIONAL QUALITY FORUM

+ + + + +

PERSON- AND FAMILY-CENTERED CARE STANDING COMMITTEE

+ + + + +

MONDAY
JULY 28, 2014

+ + + + +

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

PRESENT:
JAMES MERLINO, MD, Co-Chair
LEE PARTRIDGE, Co-Chair
KATHERINE BEVANS, PhD, University of Pennsylvania School of Medicine, Children's Hospital of Philadelphia
SAMUEL BIERNER, MD, UT Southwestern Medical Center
REBECCA BRADLEY, LCSW, HealthSouth Corporation
SHARON CROSS, LISW, The Ohio State University, Wexner Medical Center
DAWN DOWDING, PhD, RN, Visiting Nurse Service of New York
CAROL LEVINE, MA, United Hospital Fund
BRIAN LINDBERG, BSW, MMHS, Consumer Coalition for Quality Health Care
SHERRI LOEB, RN, BSN, EMMI Solutions
LISA MORRISE, MA, National Partnership for Patients
ELIZABETH MORT, MD, MPH, Massachusetts General Hospital, Massachusetts General Physicians Organization*

ESTEE NEUWIRTH, PhD, Kaiser Permanente Care Management Institute
LENARD PARISI, RN, MA, CPHQ, FNAHQ, Metropolitan Jewish Health System
DEBRA SALIBA, MD, MPH, UCLA/JH Borun Center for Gerontological Research, VA, Greater Los Angeles GRECC, RAND Health
CHRISTOPHER STILLE, MD, MPH, FAAP, University of Colorado School of Medicine, Children's Hospital Colorado
PETER THOMAS, JD, Powers Pyles Sutter & Verville P.C.
CARIN VAN ZYL, MD, FACEP, City of Hope National Medical Center

NQF STAFF:

NADINE ALLEN, Project Analyst, Strategic Partnerships
HELEN BURSTIN, MD, MPH, Chief Scientific Officer
LAURALEI DORIAN, MPH, Project Manager, Performance Measurement
KAREN JOHNSON, Senior Director, Performance Measurement
KAREN BECKMAN PACE, PhD, RN, Senior Director, Performance Measurement

ALSO PRESENT:

ERIC COLEMAN, MD, MPH, University of Colorado School of Medicine*
BARBARA CRAWLEY, Centers for Medicare & Medicaid Services (CMS)
MARY ERSEK, PhD, RN, FAAN, University of Pennsylvania School of Nursing, Veterans Health Administration PROMISE Center*
LIZ GOLDSTEIN, Centers for Medicare & Medicaid Services (CMS)
MATTHEW HASKINS, MPH, CPH, National Hospice and Palliative Care Organization*
THOMAS JAMES III, MD, AmeriHealth Mercy Family of Companies (AMFC), NQF Health Plan Council*
WILLIAM LEHRMAN, Centers for Medicare & Medicaid Services (CMS)
CAROL SPENCE, PhD, RN, National Hospice and Palliative Care Organization*
LORI TEICHMAN, Centers for Medicare & Medicaid Services (CMS)*

JOAN TENO, MD, MS, Brown University*

*** present by teleconference**

T-A-B-L-E O-F C-O-N-T-E-N-T-S

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

9:00 a.m.

MS. DORIAN: Good morning, everyone. My name is Lauralei Dorian and on behalf of all of us here at NQF we would like to welcome you to the Person- and Family-Centered Care meeting.

NQF truly does embrace the belief that patients are at the heart of everything we do and that they're our ultimate customer. So we've really been looking forward to this committee in particular and welcome you here. And we're looking forward to the next few days as we evaluate experience with care measures.

Before I have the rest of our NQF staff introduce themselves, I'd just like to go over some brief logistics. The restrooms are located out here to the right. There should be somebody at the table out there at all times if you have any questions, you needed to make a phone call or you needed a quiet office.

We have three dedicated breaks today, one at 10:45, lunch will be served provided by NQF at 12:15, and another break at 3:15.

We do have WiFi here. The username is guest and the password is NQFguest.

This call is, as all of our calls, open to the public, so please remember to mute your lines if you're on the phone. It just helps with the noise.

All of the materials are available via SharePoint. We will be screen sharing today and we have a Webinar, so please email me, Lauralei, if anybody on the call has any questions or is having any difficulty accessing the materials.

We have made reservations tonight at 6:00 p.m. at a place called Mio for dinner, which is a really good contemporary Latin American place just around the corner from here. So at lunch I'll ask if anybody's

interested and I can let them know of the specific numbers.

So with that, I will have the rest of our NQF staff introduce themselves. I'll begin with Helen Burstin, who is our chief scientific officer.

DR. BURSTIN: Good morning, everybody. Pleasure to be here with you today. We're looking forward to these meetings. These are among the most important measures in our portfolio, so really thrilled in particular to have so many consumers and patients at the table with us. So I'll have a chance to re-engage with you in a few moments as we do introductions and disclosures.

DR. PACE: I'm Karen Pace and I'm senior director on the project and will just notify you now that I'll be departing from NQF in a few weeks and Sarah Sampsel, who is here and will introduce herself in a moment, will take over for me. So should have a smooth

transition.

MS. ALLEN: Hi, I'm Nadine Allen. I'm project analyst on this project.

MS. JOHNSON: Good morning. My name is Karen Johnson. I am a senior director here at NQF.

MS. SAMPSEL: Good morning. I'm Sarah Sampsel and as Karen kindly just introduced me, I will be taking over as senior director on the project and I'm a consultant to NQF.

DR. PACE: And, everyone, these microphones, you will have to get them close. We do have people on the Webinar and also a transcriptionist, so it's important that we speak into the microphones. Thanks.

MEMBER THOMAS: My name is Peter Thomas.

MS. DORIAN: So --

MEMBER THOMAS: Oh, forgive me. I thought we were going around.

MS. DORIAN: We'll do the internal

NQF staff first. But now actually that's a perfect segue.

(Laughter)

MS. DORIAN: We're going to have -- so thank you.

MEMBER THOMAS: Didn't mean to force the issue.

MS. DORIAN: No, no, no. Helen Burstin is going to lead our disclosure of interest process. We'll do introductions with that.

DR. BURSTIN: Right, so we'll do introductions as part of this process. So part of what we try to do is try to be as open as we can about anybody's potential biases or conflicts. We recognize we can't exclude every single one of them, or we wouldn't have anybody at the table. So we are fully cognizant of that, but want to have the opportunity for all of you in particular to understand where each of you is coming from and see if you have any potential biases or

disclosures that you think would be of importance to share.

We don't need you to recite your CV; we'd be here for days, or your résumé or anything along those lines. So what we really want you to do is introduce yourself, tell us your position, where you are. And if there's anything that you think is particularly relevant to the measures before the Committee over the next two days, please, please share those.

And we'll begin with the Chairs. Jim?

CO-CHAIR PARTRIDGE: Good morning, everyone, and welcome to this Committee and to the NQF process, if you are new to it. I'm Lee Partridge and I'm going to co-chair this group over the next couple of years, which is a delight for me both because I love NQF and the work we do here, but I'm really pleased that we are starting a standing committee process in which we get to work together for

a serious length of time as opposed to being pulled together on an ad hoc basis.

I am the senior health policy advisor at the National Partnership for Women and Families, which is a 45-year-old advocacy organization here in Washington, although I am now no longer living in Washington. I'm located in Midtown, New York.

CO-CHAIR MERLINO: Thanks, Lee. I'm Jim Merlino. I'm a colorectal surgeon at the Cleveland Clinic. I'm also the chief experience officer and the associate chief of staff. And it's my first time a part of this Committee and I'm very excited because I think the work that the NQF does and the measures that it reviews and ultimately approves is really critical for improving how we deliver health care. And I think that how we take measurement to the patient and allow them to provide feedback to both drive policy, regulatory environment and also hospital operations is really critical.

And it's interesting: I was flying out here this morning thinking about what an awesome responsibility this really is. When you think about it, it is a triangle. At the top of that triangle are what we do for patients, because this really -- these measures have the potential, just as all the measures that are currently active, to really bring the voice of the patient and the voice of the family into hospital operations.

But I think we also have an ultimate responsibility that we have to ground that triangle on the other two tips, which is to look at the process in a very open objective light, make sure that we are approving things that are going to be able to be successfully operationalized across hospitals in the United States, and also approve measures that the Government on the other tip of the triangle is able to really use effectively to drive reimbursement and also improvement in health care.

So it is an awesome responsibility and I'm happy to be a part of it. And thank you for all of you, your willingness to put time in this as well.

MEMBER BIERNER: I'm Sam Bierner from the University of Texas, Southwestern Medical Center, a professor of physical medicine and rehabilitation and I don't have any conflicts of interest that I know of or financial interests that would involve here.

MEMBER CROSS: Hi, my name is Sharon Cross. I'm from Ohio State University Medical Center and I work in our Patient Experience Department overseeing our Patient and Family Advisory Council. I don't believe I have any conflict of interest.

MEMBER BRADLEY: I'm Becky Bradley. I work for HealthSouth Corporation in Birmingham, Alabama. We have 104 inpatient rehabilitation hospitals, post-acute, in 27 states, so we have multi-cultural patients and families. And they're small hospitals, but --

it's a big company with small hospitals across the nation. And I have no conflict of interest to disclose.

MEMBER SALIBA: I'm Debra Saliba. I'm an internal medicine physician who specializes in geriatrics. I work at UCLA where I'm a professor of medicine and direct the Borun Center for Gerontological Research. Our center focuses on improving quality of life for persons with long-term care needs. I also do clinical work in the Veterans Administration and research in the VA and am at the RAND Corporation.

MEMBER LOEB: I'm Sherri Loeb. I am a patient engagement strategist for EMMI Solutions, which Jim is familiar with. It's an outcome-based patient engagement company. I also have a lot of personal experience on the true need for person and family-centered care after losing my husband nine months ago and the lack of person and family-centered care and patient engagement that we saw. So

I'm thrilled to be on the Committee and hopefully we can come up with some true measures that can be implemented and can really affect the patient and be used to help them.

MEMBER MORRISE: My name is Lisa Morrise and I'm from Salt Lake City. I'm a patient and family advocate. My daughter was born 21 years ago unable to breathe or swallow, so we've had significant numbers of encounters with the system at hospitals in more than just Utah.

And I became involved first on the Patient and Family Advisory Council at our local Children's Hospital and then more and more involved nationally. And I currently serve as the patient co-lead on the National Partnership for Patients, Patient and Family Engagement Affinity Group and specialize in engaging patient advisors and developing patient and family advisory councils that are significant collaborators with their

facilities.

And I, too, am so excited about these measures. And reading them, realizing that some may need a little more tweaking, but really appreciate the efforts to engage families and patients in this way. Thank you.

MEMBER PARISI: Good morning. My name is Len Parisi. I am vice-president of Quality Management at MJHS in New York. It's a post-acute care provider. We have home health, hospice, palliative care, skilled nursing facilities and Medicare Advantage Plan.

I am very involved in a lot of these measures. I oversee the implementation and initiatives surrounding them. I'm also the president of the National Association for Health Care Quality.

MEMBER LINDBERG: Good morning. My name is Brian Lindberg. I'm with the Consumer Coalition for Quality Health Care. I and the coalition have a long history with

NQF starting back when I served on the original committee that Vice-President Gore created to create NQF. And I've served on the board and the Consumer Council off and on since.

In addition to the work I do with the Consumer Coalition, I work for small non-profits in health, aging and long-term care areas here in D.C. And before I started doing that I spent 10 years on the House and Senate Aging Committees.

MEMBER LEVINE: Hi, I'm Carol Levine. I direct the Families and Health Care Project at the United Hospital Fund in New York, and the fund is a non-profit, non-partisan health services research organization. I, too, have a lot of personal experience in this field having taken care of my disabled husband for 17 years at home without any person or family-centered involvement, and I have a granddaughter who is severely disabled.

But at the moment my biggest problem is I need some tech help to get onto the Web site. I'm technologically challenged. So anyone who can send for a person who understands why I can't get onto the Web site, any Web site, would be helpful. Thank you.

MEMBER STILLE: Good morning. My name is Chris Stille. I'm a professor of pediatrics and head of general pediatrics at the University of Colorado, also Children's Hospital-Colorado, and most of my academic work has centered around improving systems of care for children and families with special health care needs, specifically how families, primary care providers and specialty providers can work together to create the best outcomes.

And I'm happy to be here. I have no conflicts of interest and the only bias I would say I have is that anything that helps children and families I'm biased for.

MEMBER VAN ZYL: Good morning, everyone. My name is Carin van Zyl. I

originally started in emergency medicine, but came to the dark side on hospice and palliative care. And I am with City of Hope over in Los Angeles where I'm the director of quality. I have no conflicts to disclose.

MEMBER LEVINE: Oh, I didn't say that. I have no conflicts either.

MEMBER NEUWIRTH: Good morning, everyone. I'm Estee Neuwirth. I'm director of evaluation at Kaiser Permanente's Care Management Institute. I specialize in spreading both qualitative and quantitative methods for better understanding our patients and our staff to improve quality across our whole system. I'm a sociologist by training, and like probably some of you, I have some deep personal experiences. My mom died of a hospital-acquired infection about four months ago and I'm now caring for an aged parent as well. So a lot of these measures touch really close to home. And at Kaiser Permanente we're doing our best, like many of you, to spread

different ways to involve patients and families, and our group at the Care Management Institute is leading much of that work. Thank you. And I have no conflict of interest.

MEMBER DOWDING: Good morning. My name is Dawn Dowding. I'm a professor of nursing at Columbia University School of Nursing and the Visiting Nurse Service of New York, which is the largest not-for-profit home care provider in the U.S., I think still, just about. And we have home care/hospice family and nurse partnerships, behavioral health programs and we've got a big interest in looking at patient experience and how we can better improve that for our patients that we care for. And I've only been here less than a year, so be gentle with me. I'm still getting to grips with the U.S. health care system.

MEMBER BEVANS: Good morning. I'm Katherine Bevans, an assistant professor at the University of Pennsylvania's School of

Medicine and the Children's Hospital-Philadelphia. I'm a researcher. My research interest is in the development of patient and family-centered methods for integration into research. I have done quite a bit of work on the development on patient-reported outcomes for parents and especially by child-report as well. I also head up the research arm of our Family Partners Program at CHOP, which does a lot of work on quality improvement, operations of the hospital, as well as some of the research initiatives as well. I don't think I have conflicts of interest except to say that I am also biased toward children and families, as I've heard a lot of people say. So thanks.

MEMBER THOMAS: Good morning. I'm Peter Thomas and I'm with Power Pyles Sutter & Verville, which is a health care law firm right here a couple of blocks away, and I've been there since 1992.

I have 40 years of experience

walking on two artificial limbs, and that got me into health care. And when I went to law school, that's what I decided to focus on. So I do a lot of work in rehabilitation and disability issues, all focused on health care. And I don't believe I have any conflicts, but I do represent a fair amount of organizations and associations that are involved in this general space. I've never really worked on specific measures before. So I'm really glad to be part of this Committee and looking forward to working with you.

DR. BURSTIN: Great. And do we have any Committee Members on the telephone?

MS. DORIAN: We should have -- do we have, let's see, either Sherry Kaplan, Liz Mort or Kimly Blanton on the call?

(No audible response)

MS. DORIAN: Operator? Kathy, are you there?

OPERATOR: No, they're not on the call.

DR. BURSTIN: Great. Okay. So we'll return to them when they're able to join us. So thank you for those actually rather moving disclosures. They're not usually quite that moving, but thank you for that.

Just a couple of reminders. You actually represent yourselves, not your organizations who nominated you. You really are here as an individual. And the first question I'd have for you is do you have any questions of each other about anything anybody has indicated they potentially have issues with or concerns about? I think the love of all children, mothers and apple pie we're probably all good with.

(No audible response)

DR. BURSTIN: Great. And just lastly, at any time during this process if you have any concerns that there appears to be something that wasn't disclosed or a bias at any time, please come forward to NQF staff or the chairs. We really want to take the

opportunity to address those in real time rather than have them sort of fester and have issues come out later. So again, thank you for your service. Really glad this is a standing committee. I think you guys will be a great team together for the next few years.

So with that, I'll turn it back over to Lauralei.

MS. DORIAN: Great. Thank you, Helen. And thanks, everyone.

Now if you can go to the next slide, I'm just going to go over some information about your roles, the roles of the developers and some ground rules.

As Helen just mentioned, you're acting here as individual representatives for the NQF multi-stakeholder group. You will be serving either two or three-year terms, and these terms will be selected at random. We're going to be doing them tomorrow afternoon. But if you have any objections to serving a three-year term, please let me or one of my

fellow colleagues know during lunch or one of the breaks today.

You will review all of the measures. Even though we did have you assigned to specific parts of measures, we'll expect that you've reviewed all of them and will participate in the discussion today evaluating each of the measures against the criteria. You will make recommendations to the NQF Membership for endorsement, respond to comments submitted during a review period. And those can be sort of ad hoc comments that occur at any time, but we also have a dedicated public and member comment process that will occur following this meeting. And we'll go into a little bit more detail about that when we review next steps tomorrow.

You will respond to any directions from the CSAC, which is our Consensus Standards Approval Committee, and you will oversee the entire portfolio of person and family-centered care measures. So as we

touched upon, this is a big change because we are moving to standing committees. So we're looking forward to that and getting to know you.

And if you could go to the next slide, this is just some details about that change, the new function. You'll be providing input on relevant measurement frameworks. And we do have a number of similar projects that are occurring, or relevant projects that are occurring at NQF. And Karen will touch upon one of those later today.

You'll know which measures are included in the portfolio; we'll go over those in a little bit of detail later, consider issues of measure standardization, identify measurement gaps in the portfolio, be aware of other NQF activities that are related to this topic area open to any external input, and provide feedback continuously about how the portfolio should evolve.

So some ground rules for today's

meeting. We are hoping that you're prepared having reviewed the measures beforehand, and we hope that some of that extra guidance that we sent you you found helpful. You'll base your evaluations and recommendations on the evaluation criteria which we'll be reviewing. Especially when we go over the first measure we'll make sure you understand how to apply all of the criteria.

We hope that you remain engaged in the discussion, keep your comments concise and avoid dominating a discussion. You can indicate agreement without repeating what has already been said.

And one thing I forgot to note actually earlier that usually works well for us is to turn your cards on their end if you have a comment to make, and then we can make sure to call on you and to make sure everybody's heard.

And we did want to note that all Committee Members, Co-Chairs, NQF staff and

developers are responsible for ensuring that the work is completed in the time allotted, so make sure to stick to the agenda as closely as possible.

We are fortunate to have measure developers either here with us in the room or over the phone. They will be asked to briefly introduce their measures in two or three minutes before each measure discussion occurs. They have designated places at the main table. And during the measure evaluation Committee Members can offer suggestions for improvement, but we do want -- and sometimes the developers will agree and say that they'll make changes to their measures, but we do want to remind you that the suggestions are only sort of considered for future, but the votes that you're making are on the measure as currently specified.

And as is the case with Committee Members, developers, as they sit up here, can put up their cards to indicate that they wish

to respond to a comment.

So, now we're going to go over portfolio review. And I think at this point I turn it over to Karen.

DR. PACE: Okay. So the way we assign these measures; and I'll get into that a little bit, but currently we had 55 endorsed measures that were related to person and family-centered care, and 12 of the measures are actually being reviewed in this Phase 1 that you've all seen. Eleven of them were previously endorsed, so they're under endorsement maintenance review. And then there's one new measure.

Next slide. So the measures -- these are not the slides that I sent. They don't have the measure numbers. So these are the measures that are in this current project, and we'll be going through them in detail, but we have some from the acute care facility with HCAHPS. Child HCAHPS is the new one. Three-item care transition measure, inpatient

customer service, evaluation of inpatient behavioral health services, and then the clinician CAHPS.

Next slide. So we have the health plan CAHPS. We have three measures in the area of hospice and palliative care. We'll be starting with those when we start into our measure review. One is specifically about hospice care. The other two are with bereaved families that do not have -- have been in hospice care. We have two post-acute measures, the CAHPS in-center hemodialysis survey and the home health care survey measures.

Next slide. So one of the things that we wanted to discuss is that it does incorporate right now the measures in this portfolio patient-reported outcomes. And this is something that we'll be discussing a little bit further, because these measure assignments were made prior to some work that was done with our HHS partners in terms of categorizing

performance measures based on the National Quality Strategy priority and also prior to some work that we've recently done with person and family-centered care.

So I'll just mention here that patient-reported outcome refers to the concept of any report of the status of a patient's health condition that comes directly from the patient without interpretation of the patient's response by a clinician or anyone else.

In our PRO project a couple years ago we adopted this definition from the FDA, but we identified specifically four domains that we would categorize in this area of patient-reported outcomes. One is health-related quality of life, which also includes functional status; symptom and symptom burden experience with care, which is really the focus of the measures that we're looking at in Phase 1 if this project; and then health-related behaviors, such as smoking, diet,

etcetera.

Next slide. And this just makes a distinction between -- for some terminology for us to use. So patient-reported outcome really refers to the concept. What is it that we're thinking about measuring? PROM, or patient-reported outcome measure, is terminology that's often used for the patient-level scale or instrument. So in this case the patient-level survey questions or scales that are done at the patient level. And the PRO-PM or patient-reported outcome performance measure is when that patient-level data is aggregated up to look at how an organization is doing. So we may have a facility hospice-level measure, a hospital, nursing home, home health, etcetera.

An example here is some measures we have about depression. So if we think about the symptom of depression, that is something that is generally patient-reported. We have an instrument called the PHQ-9 that is

widely used, and it's a nine-item instrument that asks patients questions which they respond to in which you can quantify their level of depression. And then from the performance measure standpoint we have endorsed a performance measure that's -- so if you talk about a hospital or a clinician, the percentage of patients with a diagnosis of measured depression, an initial PHQ-9 score greater than 9 after follow-up, after treatment, actually has a decrease in their depression symptoms.

So the idea is that the concept is depression. Then we have a very specific instrument to measure a patient's level of depression and then we have a performance measure for the treating facility or clinician that aggregates the information for all their patients with depression to see how well they're achieving remission of depression symptoms.

Next slide. So experience with

care performance measures.

So, Lauralei, these are not the slides.

Okay. Before I get onto this, I'll just kind of tell you a few things that I wanted to add to this. In addition to the PRO project that I was referring to we also recently had a project about person and family-centered care and performance measure framework. So in this project the goal was to identify the basic concepts involved in person and family-centered care. The reason I want to bring this up is that in that project there was a distinction made as well as the work we've done with HHS about how to tag measures as being in the person/family-centered domain -- is that a patient-reported outcome is not necessarily an indication by itself of person and family-centered care. So this is something we're going to want to discuss with you tomorrow, but I'll give you the example.

What we've talked about is using

person and family-centered patient-reported outcome measures or tools is certainly an indication of person-centered care. So asking the person directly about their symptoms, their function is great in terms of facilitating person-centered care, but if we aggregate that say percent of patients with remission of their depression or percentage of patients that have an improved function after hip surgery, that could -- may be better an indicator of the effectiveness of their treatment.

So there's some gray area about where we should classify those measures and how they would best be -- what steering committee, what standing committee should best be reviewing those types of measures. So all of these things are -- there's kind of a gray area and we sometimes have to make some arbitrary decisions about where we place measures in terms of review for endorsement, etcetera.

Okay. There we go. Okay. That's it. So this is the definition that we came up with in the person and family-centered care project, which is an approach to the planning and delivery of care across settings and time that is centered around collaborative partnerships among individuals and their defined family and providers of care. It supports health and well-being by being consistent with, respectful of, and responsive to an individual's priorities, goals, needs and values.

Okay. Next slide. So during this project we came up with some core concepts; and I know you can't read this here, and I'll read it out to you, in terms of what are the core concepts of person and family-centered care?

So the first one is individualized care. And we kind of discussed that -- or the concept is that I work with other members of my care team so that my needs, priorities and

goals for my physical, mental, spiritual and social health guide my care.

The second concept is family. My family is supported and involved in my care as I choose. Respect, dignity and compassion are always present. Information sharing and community. There is open sharing of information with me, my family and all of the members of my care teams. Shared decision making. I am helped to understand my choices and I make decisions with my care team to the extent I want or am able. Self-management. I am prepared and supported to care for myself to the extent I am able. And access to care convenience. I can obtain care and information and reach my care team when I need and how I prefer.

So a lot of these things are definitely going to be represented in the experience with care measures, a lot of these concepts. Some of them are not, and that was part of the work of the recent measure gap

project in which we were looking at this particular framework in terms of moving forward and filling measure gaps for person and family-centered care.

So I just wanted to kind of set this up because we want to have a discussion with you tomorrow; hopefully time will allow, that we can get some of your thought for the future of where we should be thinking of assigning these measures to this standing committee. So, for example, if we have a performance measure about pain management and pain control, is that best here or is that best in a standing committee that's focused on perhaps a particular clinical topic? And so we'll have more discussions about that.

Lee?

CO-CHAIR PARTRIDGE: Karen, sometimes in the past I know it's been NQF policy to assign a measure such as the pain example you just gave to two committees. Are we trying to minimize that in the future?

DR. BURSTIN: It would be optimal to keep portfolios together, but we recognize at times you're going to want a consultation from another committee. And that's fine. So for example, our Cost and Resource Use Standing Committee just recently got consultation from our Cardiovascular Committee on all the heart-related cost measures. That's perfectly appropriate. We'd like to have measures have a home and have more of a consultative model as needed.

DR. PACE: And one of the things we may have to distinguish is a measure that has more cross-cutting applications, such as pain versus something that's a very specific -- for example, improvement in function after hip surgery might be most appropriate in the clinical topic area where they're talking about orthopedic conditions, etcetera, versus a more general functional status measure might be more appropriate in this standing committee. So we'll have to get some of your

thoughts as we're moving through this, whether there are individual measures that maybe should be moved in the future.

Okay. Next slide. So just a couple other notes about experience with care measures that were not submitted in Phase 1.

Lauralei, do you want to --

MS. DORIAN: Sure, so we have a number of measures. Some of them are up for review in Phase 2, some -- as you see on your screen, a number of CAHPS measures were not submitted to this current phase. We did have one measure, an NCQA, it was a supplemental items for CAHPS adult questionnaire that was actually withdrawn because they're currently in the process of finalizing revisions to the shared decision making and coordination of care questions that are contained within it. So you can see these other measures on your screen. I won't go through -- read out each of them.

If we go to the next slide, you

can see some more just so you're familiar with the rest of the portfolio. We have the nursing home survey, the cultural competence and literacy practices, and then the surgical care, which we're anticipating to come in Phase 2.

DR. PACE: And was there some other ones that we know are coming in Phase 2 out of this list?

MS. DORIAN: Of -- we --

DR. PACE: That young adult survey.

MS. DORIAN: The patient activation measure.

DR. PACE: No, that's not already endorsed.

MS. DORIAN: No, I don't --

DR. PACE: Okay.

MS. DORIAN: No, no, we don't.

DR. PACE: So the question will be what will happen with these measures that were not submitted for endorsement maintenance.

And that's something that NQF will have to discuss with the measure developers. Obviously, it would have to be consistent with our policies for all performance measures. Sometimes the developer withdraws them for whatever reason, as Lauralei was mentioning about 007. They could potentially be submitted in Phase 2 as potentially the surgical CAHPS maybe or potentially remove NQF endorsement. But we have to have further discussions with the developers to really identify where they're at with those particular measures. But we're just trying to make you familiar with what's in the entire portfolio.

Okay. So we've already talked about these measures that we are going to be looking at.

Okay. I'll mention just a few key points. We've talked about this on orientation in each of the work group calls, but just to kind of reinforce is that these

experience with care performance measures. The survey is a data collection method, not a type of performance measure.

NQF endorses the performance measures for accountable health care entities, not the survey or tool or instrument by itself. So obviously the survey or instrument is being used to collect the data, so it's an integral part of the endorsement process. We couldn't have the measures without the data that is collected from the patients through the survey or tool. And as I already mentioned, the performance measure actually aggregates the data for the patient served by each health care entity.

NQF endorsed performance measures are intended for use in both performance improvement and accountability applications such as public reporting and pay-for-performance. And as we've already talked about, experience with care is considered a domain of patient-reported outcomes.

Next slide. These measures, as you've all seen already, can be based on a single item or question in a survey, or multiple items or questions. For the outcome measures; and we include PRO-PMs, there's an exception to providing a summary of systematic review and grading of a body of evidence for outcome measures including these PRO-PMs. We really are asking the developers to identify and provide a rationale which could be based on evidence, and obviously that makes it stronger.

But the idea is is it something that at least one health care structure, process, intervention or service affects the patient experience that's being measured? So is there something that the hospital, the nursing home, the home health agency, the clinician can do that will really affect that patient's experience on the area being measured?

And the last bullet is something

that came out of our project in 2012 -- is that PRO-based performance measures are required to be tested at both levels, the patient-level data or score on the instrument and then the computed performance score for the health care entity. And so you've seen in the submissions or some of the additional material that was submitted after the work group calls that generally the performance measures -- first they start with testing that instrument and then as they decide how they're going to aggregate that information into a performance score there's testing of that level of the performance measure as well.

Next slide. So just a couple mentions of some of the overarching issues that we identified and some things to keep in mind for our discussion tomorrow, because after we go through these measures, and certainly with your experience of looking at them individually and on the work group calls, we really want to have some discussion with

you, some feedback and suggestions for going forward with these types of measures in the future or other PRO-based performance measures.

But specifically with the experience of care measures you've seen multiple measures submitted in one form. So we'd like you to kind of think about what works about that and what doesn't work, because in some respects it might be easier, but in other respects it's more complicated, and we just want to have some discussion with you about that.

The next one is that -- and this is not unique to these particular measures, but sometimes information was not provided in the submission form or not provided where we requested. And we know that that creates some confusion and difficulty to find the information.

Lack of identification of specific health care structures, process, interventions

or services that affect the performance measure. This relates to our kind of evidence criterion or exception for these measures where we ask the developers to identify specific things that can influence this. And so, I think this is something where in most -- in some of the cases you'll be kind of substituting your knowledge and common sense about are these things that can be affected by health care services?

One of the things that came out of the prior PRO project was that for these types of measures we should ask the developers to identify how patients were involved in identifying what's of value and meaningful to them in terms of collecting patient-reported measures, whether it's experience with care or functional status.

And so we did ask for that on the submission form. And I think generally we had little substantive information that was reported. A lot of times it was noted that

patients were involved in focus groups. And so, one of the questions I have to the steering committee to think about it is is it really necessary to ask for that? Should we just assume it? Is it sufficient? Should we just change it to a check box that they check it off? So just think about that, because again it's -- we want your feedback on improving the process. And if it's something that really doesn't add to the evaluation, then we don't want to bother developers with asking for that information.

Testing may not have been submitted for both levels. And as I mentioned, that is from our most recent guidance. And the two levels are the patient-level instrument or scale and the computed performance measure. Some of the submissions, there seemed to be some lack of standardization. Remember we're endorsing measures as a national consensus standard, and so one example here is there may have been a

statement "can choose to adjust or not adjust." And so the question is is that standardized for a national standard performance measure? If people are implementing it different ways, how would we interpret that?

And then generally there was a lack of information about case mix adjustment, development and analysis. So there may have been the final report out about which variables and coefficients, but not necessarily the development and testing process of the case mix adjustment.

Okay. Sam?

MEMBER BIERNER: I want to ask a question about that issue of case mix adjustment. I note in reviewing some of the measures they mention that specifically and some adjust for that, but across all the measures that you're following in NQF, do you have a percentage or some idea of how many use case mix adjustment?

DR. PACE: I don't have a percentage. Most of the outcome measures are risk -- for the clinical measures the term is "risk adjustment." For these measures typically it's referred to as "case mix adjustment." But the question to think about is if there's systematic differences by some patient characteristic. If it's a clinical characteristic, for example, their health status, how sick they are or how complex their needs are, if that's a systematic difference in the outcomes achieved, then it's probably something to be adjusted for. And that same basic concept can apply to experience with care measures. And I don't know the percentage of the experience of care measures that are adjusted.

Helen, did you want to add anything.

DR. BURSTIN: Last count I think about a third of our portfolio are outcome measures, and that's -- we've really pushed

and wanted more and more outcomes and less and less process. So we'll see more and more measures that have these complex case mix/risk adjustment approaches.

Okay. So where are we on our agenda? I think we're going to wait and do this when we get into our first performance measure.

MS. DORIAN: Are there any questions so far? Yes?

MEMBER BIERNER: So other than one measure all of the measures we're looking at have already been endorsed and we're looking at a maintenance review. Could you just explain exactly what that means?

DR. PACE: Yes. Thanks. We should have a slide on this, but NQF has a policy where all the measures in the portfolio are reviewed approximately on a three-year cycle. Sometimes it's a little less; sometimes more, because obviously we try to group measures in a project. So depending on

when they were initially endorsed it may have been a little less than two years or more than -- I mean, less than three years or more than three.

But the idea for that is that, first of all, in some cases the evidence may have changed for a measure, which is not as relevant for these experience with care measures. We want to know that there's still room for improvement, because if there's -- if the performance on the measure is what may be referred to as topped out, everybody's performing at 95, 96, 97, 98 percent of achieving the process or the outcome, then is it worth the effort to continue collecting the data and reporting it? So that's a key consideration for maintenance.

And then in terms of testing, the idea is that if -- especially if it -- when it was initially endorsed it may have come in with some very minimal pilot testing and we'd like to see that it really is performing as

we'd like for a performance measure. There may be other issues that come up. We have a comment period and accept comments on measures that are up for endorsement maintenance.

And then the other thing is that NQF's criteria has evolved as the quality measurement and reporting space has continued to mature, and so we want to make sure that measures that were endorsed three years ago meet the criteria as they're currently being implemented.

So, there's a variety of reasons that we would like to see updated information on all of the measures, but it is something that we're going to be looking at in the future of if there's more distinctions of what measures. But currently it's -- every measure in the portfolio is expected to go through endorsement maintenance review. It's the same criteria. They have to meet all of the same criteria.

CO-CHAIR PARTRIDGE: Actually,

Karen, I would just one other thing on that point. When you're looking at them from a maintenance review perspective, I think those of us who have been in that role on other steering committees tend to look particularly at how extensively they're used. Do we have a really lovely measure that we endorsed three or four or five years ago that nobody's picking up, and if so, why?

DR. PACE: Thanks, Lee. That is a key point. As I mentioned earlier, the measures are intended to be used for both improvement and accountability applications. And several years ago we had a committee, a task force that took another look at usability and use, and really do expect these measures to be put into accountability programs within three years of endorsement and public reporting within six years. So the question is, as Lee said, if a measure has been endorsed and not in use, the question is why?

Now, it doesn't mean that you have

to vote it down. It just means that we asked them to explain it and you need to think about that and is there a plan for getting into use? Because all kinds of things happen. And especially getting things into federal programs, for example, there can be all kinds of delays. But it is something to think about and question.

CO-CHAIR PARTRIDGE: We find many of them are so burdensome or expensive to implement that it really wasn't practical.

MEMBER BEVANS: Perhaps that's why, but I'm wondering whether there are more new measures in -- first submission measures in Phase 2. And if not, whether you have any thoughts as to why maybe there are so few new submissions. We're doing a lot of re-ups, right, but not that many new. So, thoughts about that?

DR. PACE: I think that's a good question and certainly something that as a standing committee as we kind of work together

over the next project and the next year can you have a better grasp on what's in the portfolio, what the needs are, whether there is a need for new measures. So it's certainly a -- I can't answer that question, but I think it's an appropriate topic for the standing committee to tackle.

DR. BURSTIN: And I'll just add, these are really expensive measures to develop. They're not your simple process measures. They take fairly deep pockets as well to develop them. And then they also are pretty hard to implement. So one question might be how many of these do you need? I mean, are they capturing the right patients in the right settings? That's really important.

Again, I think that's going to be a critical issue back -- and frankly, one of the reasons we went to standing committees is we'd love you guys to actually help us push the field to say these are the measures we need. How do we get them into NQF? Sometimes

they're local pilots, or local initiatives, or specialty -- or things people have come up with that are really great examples of what could be done, but they've not moved out of a local example into a national scale. And we'd love to help think about how we could bring forward really promising measures used on the ground that potentially others could benefit from if we could bring them forward.

MEMBER BEVANS: Yes, I guess my question is really more a charge for the Committee to say maybe we should do -- think about some outreach, if we really use perspective we believe really needs to be represented in quality measurement, then we need some new measures. We need to keep that flow of new measures potentially coming in, I think.

MEMBER NEUWIRTH: I was wondering if we're going to be looking at all of those 55 endorsed measures in patient and family-centered care over the course of the next two,

three years. And if so, can we get more of a global perspective on these 55 measures and how they're distributed and how they're -- across the different populations, and then also the use. And where are we as a nation in terms of using these and so on?

DR. PACE: Right. So, that was one of the slides that didn't get in. Let me just tell you in terms of the general categories. Just a second. If I can find this here. Do you have it?

Oh, okay. So of the 55 endorsed measures, 21 are in the experience with care topic area. And as you know, 11 were submitted in this Phase 1, plus we have one new. So we showed you the list of the 10 that were not submitted and that we'll have to follow up with the developers on. Twenty-one are in the kind of general area of function and health-related quality of life.

Right now there are three in symptom/symptom burden, and all three are

about pain. And then there are 10 that I just kind of put in a miscellaneous area. Some of them are about language and meeting the patient's language needs. Some are communication but from a staff survey point of view. And then culture. So this is just kind of big picture, the areas that those 55 measures are in right now that are assigned to this.

In terms of Phase 2, we're focusing on function and health-related quality of life, though we may have a few experience with care measures that end up being looked at then. And just so you know, most of the function measures; and this will come into do they belong in this portfolio or elsewhere, are clinician-assessed function measures that we'll be looking at in Phase 2. And I think there's one health-related quality of life, but I believe it's a process measure.

So anyway, I think as we look at those measures, we'll have a better sense of

whether they should stay in this portfolio or maybe they would be better elsewhere. But again, there's no kind of black and white line. Sometimes measures could be assigned a different -- more than one topic area and we just have to make an arbitrary decision, but definitely something that we'll continue to discuss with all of you as we really focus more on person and family-centered care and what measures we really want and should be in this project, in this topic.

MEMBER NEUWIRTH: So is there another committee that looks globally across these 55 measures to see where there are gaps, how there is use? So is there a way we could get some documentation about all -- you know, in a thumbnail sketch of all 55 measures so we can understand them in a more global way, but also how they're being used, and if so, who's using them and so on?

DR. PACE: Yes, I appreciate what you're saying in terms of over-viewing the

portfolio, that that would be useful, and we'll have to think about how to do that. That's one of the difficulties with these experience with care measures is that it's not 21. It's 21 times -- as you all saw, one of these submissions has 18 measures in it. And if you go by the measure title, you don't even know what it's measuring, and that's one of the reasons to have individually identified measures. And so, we need to be thinking about that.

But I think it's a good point and we definitely need to think of ways to represent the portfolios. This is kind of new work for -- or a new area for NQF in terms of starting the standing committees, but how to best implement the standing committees going forward, we'll continue to need your suggestions. And obviously being able to get a better sense of what's in the portfolio is key to you helping us manage it.

CO-CHAIR PARTRIDGE: Estee, I also

think that that's a role we can help NQF perform, because it's not easy sitting here knowing whether or not a particular hospital or community has decided to use a particular measure. There sort of isn't a national clearing house that says this measure is being used by the following entities.

So as we all -- we've got a nice patient-centered group here. We all are representing lots of different parts of the country. I think as we see some of these, or maybe even parts of these measures are being used in our own communities or our own institutions, share it.

MEMBER BRADLEY: As I was reviewing the measures that were sent to me and kind of looking across the measures, I wasn't clear. Is the assumption that all of these measures are in the public domain, or does that matter as we're reviewing? It seems like for me it may matter in terms of burden of administration, but I didn't see that

referenced. I think it may be in one of the measures that I looked at.

DR. PACE: Yes, all these measures are in the public domain. NQF does have a policy where we can accept and endorse proprietary measures, but that would be just -- there's a whole process of what needs to be disclosed for that. But in this case all of these measures, the stewards, many of them are developed and supported by federal agencies, so those are public. And then the ones that were developed by private organizations have agreed to make them publicly available.

MS. DORIAN: Are there any other questions about the process or anything else we discussed?

DR. PACE: So why don't we go on to the --

MS. DORIAN: Because we're actually right on time, within one minute, to move on to our consideration of our first measure. And so, I think before we actually

do that we have -- we wanted to refresh your memories with a review of the measure evaluation.

DR. PACE: Right. Would you go back to the slides, because I want to just do a couple and then we'll -- so, and let me just -- those of you who were having technical problems, at our break our technical people are going to come up. Oh. Fixed? Okay. Thank you. So everybody's good? Wonderful.

Okay. So for those of you who are going to be looking at the measure, we're going to be starting with 0208, but what the process we're going to follow is we will start with a brief introduction by the developer. I will guide discussion going through our criteria much like we did on the work group calls. We'll ask the assigned discussants to comment on how the measures do or do not meet the criteria.

Then we can have full committee discussion and questions. And then we will

vote on the criterion before moving onto the next one. So, we'll talk about evidence, we'll see if there are any issues, and then we'll ask you to vote.

Because there are multiple measures in one submission form for most of these, we'll vote for all, but I'll say that any could be pulled out to be voted on separately. So I know on some of the work group calls there were maybe particular measures that people questioned or there may have been issues with, so if it's something like that that you think requires -- that you would vote differently on one of the 12 measures, then certainly raise that so we can pull that out to vote on separately.

Okay. Next slide. So we're going to start with looking at the evidence criterion. And basically the reason we are concerned about evidence is this is the foundation of using as a quality indicator for validity. And for process and structures we

want to make sure it is something that all health care units should be implementing. So again, we're implementing national standard performance measures.

But as I mentioned earlier, for outcomes and including these patient-report outcomes of experience with care we really are just looking for identification and rationale that there is something that the health care unit that's being measured and evaluated can influence.

So now let's look at our first measure, importance. And we'll start with evidence for this family evaluation of hospice care.

And before we do that, Lauralei, who's going to introduce the measure?

MS. DORIAN: Oh, yes. Do we have Carol Spence on the call, or anybody else from the National Hospice and Palliative Care Organization who would like to introduce this measure?

DR. SPENCE: This is Carol.

MS. DORIAN: Great. Thanks, Carol. Go ahead.

DR. PACE: Okay. Thank you. So the Family Evaluation of Hospice Care, or FEHC, measure has been -- we -- this is up for endorsement. So the survey itself has actually been in the hospice field since 2003. And National Hospice and Palliative Care Organization, NHPCO, has also -- goes out to hospices, and it has been widely adopted. We have over 2,000 hospices currently using this survey and getting reporting from us that includes this measure.

We feel that this is a caregiver post-death survey, so it is designed to get the caregiver's perception of the quality of care that was delivered. And we do a post-death survey because that gives -- has the advantage of providing the overview of the full episode of care. Also, both the patient and the family are the unit of care for

hospice, so this is not a proxy survey per se. In the 17 items that are included in the measure there are some which include the caregiver's perception of symptom management, for example, but many of these questions are also directed at the caregiver's direct experience with the care provided by the hospice to them, as well as to the patient.

And then just one other brief comment. In the material that Karen so assiduously provided in the review of the measure, she did note that CMS has developed a hospice survey, and just a couple words on that.

For several years CMS, in their rulemaking had mentioned family evaluation of care and our survey and measure and had considered using it, but then they decided to go ahead and develop a Hospice CAHPS survey that would join the family of CMS CAHPS surveys. And Dr. Joan Teno was part of that development, and we worked over all these

years with Dr. Teno on our measure development. So the Hospice CAHPS survey borrows heavily from FEHC. They have a lot in common.

We decided to go ahead and to go for endorsement maintenance on this for a couple reasons. First of all, the Hospice CAHPS survey is not going to be required across the board for implementation until April 2015 and we felt that hospices still should have an endorsed survey to use up until the time that they need to switch to that.

The other reason is that hospices -- not every single hospice will be using the CAHPS survey. There are eligibility requirements for that. So a hospice that has fewer than 50 eligibles deaths will not be participating -- for the previous year will not be participating in that Hospice CAHPS survey. And we're not sure exactly what that proportion will look like at this point, but one of the eligibility criteria for Hospice

CAHPS is going to be that the patient had received hospice services for at least 48 hours.

And we do know that a third of hospice patients die within seven days or less. I'm not sure what the statistic is exactly on that 48 hours, but there will be smaller hospices out there who will not be participating in the Hospice CAHPS. And we wanted them to continue to have the availability of the FEHC survey if they so chose to use it, because this measure has definitely become very much part of hospice's QAPI programs and we wanted them to have the ability to continue to use it. So, thank you.

DR. PACE: Okay. Yes, Brian?

MEMBER LINDBERG: This is Brian Lindberg. I just had a clarifying question related to that. How will one determine which of the potential two surveys that an individual family would complete?

DR. SPENCE: Well, a hospice that

meets those eligibility requirements of having more than the 50 deaths is going to have to use the Hospice CAHPS survey. So it would be again up to the hospices that don't meet those eligibility requirements to decide whether they want to continue to use FEHC perhaps, since they're already using it, or switch over to the Hospice CAHPS survey. But the public reporting and the submission is only going to be done by those hospices with 50 or more deaths.

CO-CHAIR PARTRIDGE: Sherri?

MEMBER LOEB: This may be in here; this is Sherri Loeb, and I apologize if it is. Is this for any hospice care? It does not need to be an inpatient hospice facility? So it can be home hospice?

DR. SPENCE: It's all deaths.

MEMBER LOEB: Okay.

DR. SPENCE: It doesn't matter what the setting or where the place of death is. It doesn't matter. As you may know

patients sometimes change settings of care. Hospice is an approach to care as opposed to setting-specific. So it can be inpatient, hospice inpatient. It can be someone who had the majority of their care at home. So it covers the entire episode of hospice care regardless of setting.

CO-CHAIR PARTRIDGE: Thank you. Debra?

MEMBER SALIBA: Just a couple of questions. I'm Deb Saliba. One of the questions is how many hospices are there in the U.S.?

DR. SPENCE: That depends on how you count hospices.

MEMBER SALIBA: Okay.

DR. SPENCE: The ones that are Medicare-certified the way CMS counts them is by their CMM number, but -- or the Medicare provider number, but hospices can have more than one site to bill under that number. So when we count them, we count them by site or

location rather than by parent. So we have -- our number comes up to around a little over 4,000. CMS has just over 3,000. So again, it depends on how you want to define a hospice.

MEMBER SALIBA: Because I noticed that somewhere in these materials that the majority of the participants were not-for-profits in the survey.

DR. SPENCE: I don't know that that is necessarily the case. It could be 50/50, but I don't -- I can't off the top of my head -- Matthew Haskins, our analyst, is also on the line.

Matt, do you have a number for that?

MR. HASKINS: Yes, so the utilization rates -- she's right, at the site level it's a little more strongly associated with -- not strongly associated, but there is a little bit more not-for-profit providers utilizing this survey at this point in time. But, yes.

MEMBER SALIBA: Okay. So --

MR. HASKINS: And again, we're talking about the site level though. We're not necessarily talking about the agency level.

MEMBER SALIBA: Ah, okay.

MEMBER BIERNER: It looks like it says 746 not-for-profit in the specifications from the developer.

MEMBER SALIBA: Yes, which is I think a slightly higher proportion than what there are nationally.

DR. PACE: So let's walk through the criteria and then we can -- where this may be relevant can discuss that.

So, our first criterion is about the evidence. And basically what again what we are looking for here, are there health care actions that can influence this? This particular survey has produced -- they're submitting one performance measure that really

accommodates all of the major topics in the survey. So, and they provided -- Nadine, will you go to the -- they provided information about health care actions that can influence this particular experience with care.

So, and again, on the agenda that you all have at your seat we reminded you who was reviewing different aspects of this measure, so those of you who were specifically reviewing importance, the feasibility and use and usability, that group, and anyone else -- but if you have any comments about or any issues with the fact that the hospice unit can influence experience with care on this measure.

CO-CHAIR PARTRIDGE: Those can be either pro or con.

DR. PACE: Right.

CO-CHAIR MERLINO: I'm on the list, so I would add that the evidence to support process improvement is pretty strong with this survey.

DR. PACE: Any other comments?

(No audible response)

DR. PACE: So I think we will then proceed to voting on this. And, we'll --

MS. DORIAN: Let me just check before we do that to see if we have steering committee members on the phone at this point.

(No audible response)

MS. DORIAN: Okay. So you all should have received -- just raise your hand if you don't have a voting device. Oh, you don't have one?

MEMBER SALIBA: I just have a question.

DR. SPENCE: Oh, sure.

MEMBER SALIBA: Are we going to talk about the reliability and feasibility? Are we just voting? We're voting section by section?

DR. PACE: Yes. Right. Criterion by -- it must pass criterion by must-past criterion. So this is strictly about the

evidence criterion, which on these measures is are there health care actions that the hospice can take that will influence the experience?

MEMBER SALIBA: Okay.

MS. DORIAN: So, Nadine now will review how we're going to be using these voting clickers -- devices.

MS. ALLEN: Please point the clicker towards my direction. This will be the laptop that will be capturing all your votes. When you click for your vote, for example for evidence 1A, you can select 1 for yes, 2 for no. You do not have to hit send on the clicker. Just hit the number and it will register. You can change your vote at any time. You have 60 seconds to vote. And once the 60 seconds is up, we capture that vote. Any questions?

DR. PACE: And the voting slides you'll see on the two screens at the end. So in this case we're talking about that there was a rationale that supports the relationship

of the health care outcome or PRO to at least one health care structure process intervention or service.

MEMBER VAN ZYL: You mentioned that the work sheets from the group calls would be available, but I cannot find them. Is there any way that we can pull them up just so that the rest of the Committee Members who weren't on the work group know what the comments about this particular section were?

DR. PACE: Yes, we'll have someone come down and help you --

MEMBER VAN ZYL: Okay.

DR. PACE: -- pull them up.

MS. DORIAN: Somebody should be coming up soon.

CO-CHAIR MERLINO: Len?

MEMBER PARISI: I had a question about what you just said, Karen, that the criteria would be that it would be actionable based on the performance measure, so that would be the composite score in this case?

DR. PACE: Okay. So in this case the composite is made up of all of these elements. And so, yes, that there is at least one health care structure, process, intervention or service that can influence the overall experience that's captured by this particular composite.

MEMBER PARISI: Would it be -- and my understanding is the composite score in itself would not be actionable. You'd have to do a drill down. So is that what you're -- also mean?

DR. PACE: Yes, that -- so it's not that the composite score will tell you what to do, but there are things that you -- based on what's included in that composite score you can drill down and look at your own care and find out what you can do. And that's not just an issue with experience with care PROs, but with outcome measures in general. So say you were doing a measure of mortality. That outcome measure itself doesn't tell the

hospital or home health agency what they have to do. They would have to drill down and look at their care processes to see if their patients who die versus the patients who survive -- there are different things that they should be doing.

So it's really about -- not so much that they would have to drill down, but are the things that are being captured in this composite actionable? Are there at least some things that will be influenced by the hospice?

MEMBER PARISI: Just one last question. The issue of the language, the translation of the instrument, I had raised that during our work group call, and where does that figure into this? According to the documentation here it says it's only in English. I know that CAHPS will be required to be in multiple languages, so there will be some discrepancy there.

DR. PACE: It's a good point. Yes, let me think about that. It won't be

here in evidence though, but I'll give that some thought.

And just one comment about the fact that there will be a Hospice CAHPS. It really shouldn't influence how we vote on this, how you evaluate this measure. In the future when Hospice CAHPS comes to this Committee for endorsement, then you'll need to be thinking about that would be a competing hospice to this one and which one would be the better one for NQF to endorse?

But I think that -- the language issue I believe we should probably think about under validity, because if it's knocking out a big portion of the population -- we'll hold that for now.

MEMBER MORRISE: I was wondering; and maybe the person who is on the telephone could share, was this a measure only for family members of patients who were 18-plus, or did it include pediatric patients who may have been in hospice care and their bereaved

family members?

DR. SPENCE: This is Carol. It does not include pediatric, you are correct. This is for patients who are 18 years and older.

CO-CHAIR MERLINO: Debra?

MEMBER SALIBA: So, just two questions: Can this be endorsed in a time-limited way? Can we say that it's endorsed through 2015 or 2016 so that it could be reconsidered with the CAHPS hospice?

DR. PACE: No, we don't do time-limited endorsements.

MEMBER SALIBA: Okay.

DR. PACE: So it would be endorsed until such time that a Hospice CAHPS would be brought to NQF to be considered for endorsement.

DR. BURSTIN: So at that time the measure would be endorsed for three years, but at the time a new measure comes forward that's competing against it, they will have an

opportunity to look at them together.

MEMBER SALIBA: Effectively it is sort of time -- when another competing measure comes up?

DR. BURSTIN: We don't use that term just because it is truly endorsed for a three-year term, but we will make sure there's an opportunity to look at them side-by-side as the new measure comes forward.

MEMBER SALIBA: So the other question that I had was -- gets to the issue of actionability. I thought it was somewhat striking that the average composite scores over the past three years have really not increased. I mean, it's a statistically significant increase, but it's basically the same. It goes from 85.37 percent to 85.51 percent. It's an average. So it could be that there were some facilities that improved and some that didn't.

I don't know if the developers have some insights about anything breaking

down more at the organizational level as opposed to this mean, because this would tell me that it would not be that actionable because we're not seeing the scores move very much over three years.

MR. HASKINS: I can comment on that; this is Matthew Haskins with NHPCO, if you want me to at this stage. Or if you want me to hold it to another stage I can as well.

DR. PACE: Go ahead, Matthew.

MR. HASKINS: Okay. So we've been internally looking at that number recently, actually before NQF submission started, but what you're seeing there is a lot of organizations coming in to and out of the use of FEHC. It's actually more coming into the use of FEHC. So the average change gets diluted when you get all these folks that have started using FEHC for the first time.

When you look at organizations that have utilized FEHC continuously, over a period of time you actually start to see more

improvement on those FEHC scores.

MEMBER SALIBA: Thank you. That's really helpful.

DR. SPENCE: Thank you.

MR. HASKINS: You're welcome.

MEMBER THOMAS: I have two quick questions, hopefully. The first is that if we're going to be voting on each one of these very individually, there's -- maybe not with respect to this particular vote we're about to take, but there are a number of instances that I reviewed very much depth where the data won't be in the right place, it won't be submitted, but it might be in some attachment or something and that's noted. The case may not have been well made and yet it still is viewed as highly or moderately appropriate kind of evidence. And so I'm grappling with what to do with that, number one.

And number two, what are the consequences of this vote? A, I assume this is majority rule, but I don't know that. And,

B, if there is a failure of one of these must-pass elements, then what actually happens to that measure? What's the consequence?

DR. PACE: So, Nadine, will you go back to the beginning of the slides so we can review? No, go back to the -- okay.

So good question. So thank you for kind of making us pause here. So the must-pass criteria, which are all of those under importance: evidence, performance gap, priority and then reliability and validity -- each one of those is a must-pass criterion, meaning if it doesn't pass, then we stop. It doesn't go forward. However, this year it's not just a simple majority. We've instituted a process where if it's in the 40 to 60 percent range, then we continue on because we consider that not sufficient consensus to say it should stop or affirmatively move forward.

So we will be during the voting identifying the percentage and anything that's in the 40 to 60 percent range you'll continue

to vote. If it were less than 40 percent that says it met the criteria, it would stop. So that's kind of where we're at with the voting.

Now, in --

CO-CHAIR PARTRIDGE: Karen?

DR. PACE: Yes, go ahead.

CO-CHAIR PARTRIDGE: Forty to sixty percent of those of us who are at the meeting or absolute number?

DR. PACE: Right. We need a 75 percent quorum, so we are right at 75 percent.

MS. DORIAN: Right. And we don't actually currently have anybody on the phone right now who will be voting, so it is limited to the 17 who are in the room.

CO-CHAIR PARTRIDGE: Okay.

DR. PACE: So good questions. I'm glad you're bringing these up. And I should also mention that it's -- in terms of passing, a moderate or high get combined. So moderate is passing. High is passing. Low is not. And insufficient information to determine

whether it meets the criteria also would be not passing.

So your next question is if you think they haven't made the case, there's a couple things: One is if it's clearly that it -- let's take reliability. If what they submitted demonstrates it's not reliable enough, then you would vote low. If you feel that they haven't submitted the information for you to tell whether it's sufficient reliability or not, then you would vote insufficient. Essentially, they have the same effect, but it's more clear what the issue is.

And if it's insufficient information, for example, the developer might, during the comment, period be able to provide additional information or tell us about it today on the phone. That would kind of be able to fill that gap so that when you re-look at these after the comment period that it might change what you thought about it.

Katherine?

MEMBER BEVANS: This is a question for the developer. Are we ready to get back to that or --

DR. PACE: Is it about evidence?

MEMBER BEVANS: It is about evidence.

DR. PACE: Okay.

MEMBER BEVANS: With regard to the importance of the five dimensions that are identified for quality, could you comment on how if at all patients or care providers were involved in the selection of that content or the prioritization of that content?

DR. SPENCE: You're talking about the domains that the group --

MEMBER BEVANS: Yes.

DR. SPENCE: Those were not a priori. Those were based on factor analysis.

MEMBER BEVANS: Well, when you originally developed the measure you generated a number of items. How was the content of those items determined initially?

DR. TENO: Oh, so maybe I --

DR. SPENCE: Go ahead, Joan.

DR. TENO: -- can answer.

DR. SPENCE: Go ahead.

DR. TENO: Yes. Hi, Joan Teno. I'm a little bit jet lagged because I just got back from California late last night, but we assembled an expert panel. We did focus groups with bereaved family members that have a published paper and we published a conceptional model. Prior to that we also worked with Anita Stewart to develop an overall conception model that's based on a Donabedian model. So we started out with an overall conception model based on a Donabedian model. Then we did expert panel, focus groups and then took the items back to the expert panel and had them do a ranking thing which was under the control of the hospice. And much of this has been published in JPSM.

In development of the second version instrument, when they looked at the

core with the expert panel again, we had them do another ranking procedure where they said the degree to which it was under their control, it's important. We also just recently have completed and have our first article published on focus groups with bereaved family members again. And again, we did a total of 16 focus groups from 5 regions across the country. And again, they reaffirmed sort of this core contents area.

DR. PACE: Okay. And let me just say that that's a question that we address under high priority, how the patients were involved. So right now we're voting on the question whether there are health care interventions and actions that can influence the patient's hospice experience as captured in this measure.

CO-CHAIR PARTRIDGE: And we are about to vote. Everybody ready?

(No audible response)

CO-CHAIR PARTRIDGE: Okay.

DR. PACE: Wait a minute. Nadine, you need to go to the right voting slide. Okay.

All right. So one is yes, two is no.

Oh, so, let me just wait until -- you'll see in the lower right-hand corner when the timer starts. It will count down. So just give us a moment.

MS. DORIAN: So you can begin voting now.

(Voting)

That's 16 yes, and 1 no.

DR. PACE: Okay. Let's move on to then the next question, which is performance gap. And here we want you to discuss does the performance data provide a demonstrated gap in care, which can be variability or overall less than optimal performance to warrant a national performance measure? Does the performance data provided demonstrate disparities for certain population sub-groups?

So this is where we ask them to provide some information about performance on the measure, and it can -- a performance gap can either be on the facility performance or disparities in care that would warrant us -- that there's opportunities for improvement.

So for those of you who looked at importance, Carol, Brian, Sherri Loeb, Jim and Carin, any thoughts about performance gap?

MEMBER LOEB: I definitely think there's a performance gap and I think this is just crucial to helping to eliminate that performance gap having some standardization.

CO-CHAIR PARTRIDGE: And this is Lee. I would just add that I thought that the answer to Debra's earlier question about the trends and fact that the more -- that those that have been using this survey longer do demonstrate some improvement. But the newer users are discovering that they have issues to address probably also reinforces Sherri's position.

Are we ready to vote?

DR. PACE: Right. And let me just mention, those of you who may not have taken the deeper dive on these measures, if you have questions for your colleagues that did, definitely bring them up as well.

Oh, on this particular -- we ask you to rate high, moderate, low or insufficient information to rate it. So it may be that if it didn't look like there was wide gap, you might want to vote moderate. If it looked like a large gap, high, and so forth. So any questions about voting before we start?

(No audible response)

MS. ALLEN: So one is high, two moderate, three low and four insufficient. We're ready to start voting.

So 6 high, 11 moderate, no low and no insufficient.

DR. PACE: Okay. So now we'll move onto 1C, which is high priority. And I

think the key issue here obviously is that it's an important area in terms of numbers of patients and consequences of quality, but the thing that's specific -- the other issue specific to PRO performance measures is how it was determined that the target population values each individual measure focus and finds it meaningful.

So in this case I think we heard a little bit about that.

MEMBER BEVANS: Just a follow-up question. Sorry to pose that question at the wrong time before, but in the identification of these concepts you mentioned your expert panel. Could you describe the panel? Who was that made up of?

DR. PACE: Carol or Joan, could you describe the focus groups?

DR. SPENCE: I mean, I can answer, but Joan actually did that work, so it would probably be better if she -- Joan, are you there? Can you --

DR. TENO: Yes. No, I'm sorry. I'm just muting myself and my iPhone thumb was not doing well, so I had to put in the number.

So we included experts from nursing, medicine, cancer care, people who were administrators in local hospices. And then in the original sample, industries and sample we also -- in addition to those experts we also included five persons who were the quality managers of their local hospice quality --

DR. PACE: I think we're most interested in the patients and families, or the families.

DR. TENO: Oh, the patients and families? Okay. I'm sorry. I thought you were talking about the expert. I thought the question I heard was about the experts.

So patients and families. So the original focus groups were done throughout New England dealing with bereaved family members from hospice, from people who died in a

nursing home, and people who died in an acute-care hospital and people who died at home without services.

In the subsequent follow-up papers that we're going to be publishing, one is already in press, we did -- we included bereaved family members from hospice, from six hospice programs. We chose two of those hospice programs specifically because they had a high Hispanic population and one had a high African-American population. And then one hospice program was located in a rural location. So we tried to get a geographically diverse group of family members. And we also tried to do separate focus groups that were dependent on the levels of care. So we focused on people who died in the hospice inpatient unit, people who died in a nursing home or assisted living facility and people who died at home.

Does that answer your question?

MEMBER BEVANS: It does. A quick

process question: Unless I'm missing it, I don't think this information is in the submitted material and I'm -- should we consider it in our voting of the sufficiency of the evidence? Okay. Thank you.

DR. PACE: Yes, I think in this case -- I think it's a good point and we did ask for it. And that was one of the things I brought up earlier. We got minimal information in most of the submissions. And so, it's something for you all to think about for the future, whether it's something we need to press harder to get or whether it's not really that useful for your evaluation.

CO-CHAIR PARTRIDGE: Debra?

MEMBER SALIBA: As one of the deep divers on this section, Joan, this is Deb Saliba. Welcome back. We changed coasts. I wanted to ask you a question about the data that you're talking about right now. Is this for the care instrument, or is this for the --

DR. TENO: No, this is -- so to

explain a complicated story, there has been two processes that have been ongoing that I'm involved in. I have an RO1 and during that RO1 we looked at the core of the FEHC again, the Family Evaluation of Hospice Care, and then we looked for additional items. And that is the project that we just did the 18 focus groups on. The first paper has been published from that. And then we're -- I'm currently summarizing all those focus groups for my report from our RO1.

Also, at the same time we've been working on the parallel process of the Hospice CAHPS tool as well. I've been involved in that. So I've overlapped on sort of both projects. And it's probably my fault just given -- I've had somewhat of a crazy spring in that my dad had a 22-day ICU stay where he almost died and I probably just have not been able to attend to all these issues when I've been dealing with a family emergency.

MEMBER SALIBA: I'm sorry to hear

about that, Joan. I hope he's okay.

So I guess --

DR. TENO: He finally got discharged yesterday. So, yes.

MEMBER SALIBA: Yay. Congrats.

DR. TENO: So, thank God.

MEMBER SALIBA: So my question is about the care versus the FEHC. Not the CAHPS. The care. So the one that actually goes across multiple settings --

DR. TENO: Yes.

MEMBER SALIBA: -- that you developed and published the paper in JAMA about. And I know the FEHC is derivative of the care, but I guess I'm sort of looking at the care as one that does multiple settings and would allow us to cross --

DR. TENO: Yes, right.

MEMBER SALIBA: I know we don't want to skip ahead, but in a way I think it's an important question about the importance, because this is only looking at hospice versus

hospice, whereas the fundamental question may be what goes on in different care settings with and without hospice?

CO-CHAIR PARTRIDGE: But we do have to focus right now on this one. We understand --

MEMBER SALIBA: Well, I think it's an important issue then that this is only limited to voluntary participation --

CO-CHAIR PARTRIDGE: Yes.

MEMBER SALIBA: -- of hospices as opposed to the fact that people often get end-of-life care in multiple other organizations and settings.

CO-CHAIR PARTRIDGE: Yes.

MEMBER SALIBA: So I think for importance that's a limiting factor for this particular one.

DR. PACE: So, yes, our priority criterion is about the topic area. That gets into kind of the specific measures and what patients are included. So this is really

about is it a high-priority area and were patients and -- or families involved in identifying what should be measured. And then under reliability and validity the specifications is where we get into the specific measure, so in this case limited to hospice patients. But it's a good point.

And then if we have time, we'll start talking about competing measures, which we have three in this space: the care, this one and the VA measure where we can talk about some of the issues across them.

CO-CHAIR PARTRIDGE: I'm concerned about keeping us on the clock, and this our first measure. I think it's important to discuss a lot of these issues, but Jim and I are going to be a little mean as the two days go on.

(Laughter)

DR. PACE: Okay. Any other questions about high priority?

(No audible response)

MS. ALLEN: So we're going to start voting on high priority. One meaning high; two, moderate; three, low; and four, insufficient. Starting the clock now.

(Voting)

MS. ALLEN: One more. Please vote again.

(Voting)

MS. ALLEN: Perfect. Thank you.

And the results are: 11 high, 5 moderate, 1 low, and no insufficient.

DR. PACE: Okay. We're going to move onto reliability and validity. So before we actually talk about this particular measure, I'm going to do a couple slides about reliability. We'll do that first and then -- before we go to validity.

So, Nadine, can you pull those up again?

Okay. So just a couple of points about reliability and validity in general. Why are we concerned about it? Again, NQF-

endorsed measures are intended for use in accountability applications such as public reporting and pay-for-performance. And the scores are used to make conclusions about a health care unit's quality. It could be patients and consumers making conclusions, health plans, purchasers and the health care units themselves. And if it's not an indicator of quality, it's a validity issue. And if there's too much measurement error, it becomes a reliability issue.

So, next slide. A few notes on reliability. And we had this in your briefing memo, but reliability is a way to quantify the amount of random measurement error and how well we can confidently distinguish between patients when we're talking about patient-level measures or between the health care units when using the computed performance score.

Low reliability means that it could be difficult to distinguish between

patients or between the health care units and could lead to misclassification. So, and then reliability is not a fixed property. It can vary depending on the groups included in the measurement, the patients and health care units.

Next slide. Just a couple notes. I think most of you have noticed this, that common testing approach for patient-level scales or measures is internal consistency reliability which assesses the consistency of responses for items in the scale. This is often measured by Cronbach's alpha, which is the average intercorrelation among the items, and the values range between zero and one.

A common testing approach for single item measures is test-retest which is measured by interclass correlation coefficient. Again, values range between zero and one for that. Or it could be a Pearson correlation. And I don't know that we've seen any test-retest in the submissions that we

received. I don't believe so.

And then common testing approach for a performance score is signal-to-noise analysis which reflects the proportion of total variance that is due to the signal or the real differences between units. It is often measured by ICC or inter-unit reliability which is based on the F test from INOVA. So basically the values will range between zero and one. And rule of thumb is that a minimum reliability of 0.7 to 0.75 is needed when you're trying to make distinctions between various units.

So with that, we'll get into discussion of the reliability information that was submitted for the Family Evaluation of Hospice Care. Also, the other thing that we include under reliability is that the specifications are precise. So I know everyone looked at specifications, so if there were any questions about the actual measure specifications, whether there were any --

anything that was unclear, this would also be the place to bring that up if you had any questions.

But for this, for looking at the reliability and validity, Katherine, Don, Lenard, Deb Saliba, if you have anything to start us off.

CO-CHAIR MERLINO: I would just add. So one thing that Lee and I asked last week, or two weeks ago, was that the NQF staff really drill into reliability and validity on each one of the measures recognizing that, with the exception of a couple of people around the room, we're not statisticians. Unless you use this all the time, it can be a little tricky. So the summaries that they passed out really get into the issue of whether it's appropriate or not and whether the test results are valid.

MEMBER DOWDING: Yes, this is probably more to do with validity than reliability, but it's picking up on the point

of the lack of translation into Spanish. If you look at the people who've responded to the survey, 97 percent are white. And I think there's an issue about whether or not the questions that are being asked are capturing diversity, because we're obviously getting a vast proportion of people who are choosing to reply and respond who are from a certain ethnic group.

DR. PACE: Can you hold that validity, because I think that is -- and I know it came up in one of the prior questions as well and I think it's something worth noting there and discussing.

CO-CHAIR PARTRIDGE: Any other comments from either the group that took the deeper dive or the rest of us?

MR. HASKINS: This is Matthew Haskins from NHPCO. I wanted to jump in on that last comment about the population demographic mix.

CO-CHAIR PARTRIDGE: Can we hold

that until we get to that one, please?

MR. HASKINS: Oh, I thought she brought that up in the context of validity.

DR. PACE: We're trying to focus on reliability right now, so we ask to hold that until we talk about validity, please.

MR. HASKINS: Okay.

DR. PACE: Thank you.

MR. HASKINS: No problem.

CO-CHAIR PARTRIDGE: Ready to vote?

DR. PACE: So does anyone else have any other comments about this? They did test at the patient level instrument level with Cronbach's alpha and the statistics were good. And then they did the signal-to-noise based on INOVA and IUR, and it was 0.76 for this composite. So based on our algorithms; and those are also at your place, this could be rated high on reliability unless you have any concerns or issues. If you didn't think the sample was large enough. It was done in

1207 hospices. Or if you had concerns about the testing. But any questions before we move on to voting on reliability?

(No audible response)

DR. PACE: Will you bring up the reliability vote?

MS. DORIAN: And I'll again just confirm, do we have any Committee Members on the phone? Liz, are you there?

MEMBER MORT: I am here, yes.

MS. DORIAN: Well, great. Welcome. Would you like to briefly introduce yourself since you've just joined? We went around earlier this morning. So if you just wanted to say a few words about your background?

MEMBER MORT: Yes, I apologize for signing in late this morning. My name is Liz Mort. I'm an internist at Mass General, senior vice-president of Quality and Safety and chief quality officer, and I've been working in quality measurement for a couple of

decades and delighted to be part of this panel. I apologize I could not be there in person today.

DR. PACE: Great. Thank you. And did you feel comfortable voting on this measure via the chat function in the Webinar, or would you prefer -- I'm not sure when you joined, if you wanted to start voting in the next --

MEMBER MORT: I'd rather start at the next go-round because I came in partially through the beginning -- I didn't start at the beginning of this measure's presentation. I'd prefer to wait until the next one, if that okay with you.

DR. PACE: Okay. No, that's perfect. Thank you. I'll send you a message in the Webinar chat to make sure it's working and you can receive it.

MEMBER MORT: Thanks.

DR. PACE: Okay. Nadine? Any other questions about voting on reliability

for this measure?

(No audible response)

DR. PACE: Okay.

MS. ALLEN: So now we're voting on reliability. Same answers: One, high; two, moderate; three, low; and four, insufficient. Now beginning the vote.

(Voting)

MS. ALLEN: Perfect. Thank you. Results are as follows: Thirteen high, four moderate, no low, no insufficient.

DR. PACE: Okay. So now we'll move onto validity. And I have a couple more slides, Nadine, to talk about some notes on validity.

MS. DORIAN: Yes, they're on there now actually. It's in the in-person meeting folder. There are PDFs in there.

DR. PACE: And we should just mention that the notes that Jim was referencing that staff put together, I know Lauralei sent out notes last week, is in -- at

the beginning at the measures. It says "Staff Review." There's an Excel file that has each of them.

Okay. So just a couple notes on validity. This is primarily about the degree of confidence and the inferences or conclusions that are made based on the basis of the measure, either about the patients at the patient-level scale or about the quality of the health care units when you use the computed performance score.

Again, this is not a fixed property and generally is developed over time. I think it's important to realize there are many approaches to validity testing and often with various names. So you'll see this, and it seems like different disciplines use different terminology, but it basically comes down to testing some hypothesis about how the measure performs. For example, how it relates to other measures of the same construct, how it's related or different from other

constructs, how it distinguishes groups of patients that are known to differ on the construct, or predicts performance on another measure.

CO-CHAIR MERLINO: I think there's been an question of race and language inclusion. Maybe we could just have the developer briefly comment on that?

MR. HASKINS: So, this is Matthew Haskins again. Regarding the concerns related to having it being an English version, when you look at the distribution of populations on the graphics in our test sample, those populations on the graphics are very, very similar to the population demographics that are utilized in hospice across the nation at this point in time. So while it is a limitation of the survey to not have it tested -- and we do have untested Spanish versions and other language versions that are being utilized, that -- well, on the tested version that is in a different language, the English

language version is still capturing those -- not those values, but very specifically the same patient demographics that you see utilizing hospice across the country.

DR. TENO: Matthew, can I just add a little bit of a correction to that?

MR. HASKINS: Sure.

DR. TENO: Because we did make a Spanish version available that we did test.

(Simultaneous speaking)

DR. TENO: The issue is despite making it available, no one really felt the need to use it.

DR. SPENCE: Joan, we do have some hospices that pick it up. We've given it out on request. So far the feedback we've then got back from it was that they felt again that there was some particular wording that they said, no, this is not what our particular population -- one of our translators wants to change.

So what we did was we allowed

those hospices that were using Joan's tested version to make the changes they felt were necessary. However, we asked them to keep the questions in exactly the same order, not to delete any. And then when they came to give us their responses, which was done online, that they then are able to parallel, go ahead and choose the responses based on the English version because the formatting was identical.

So within what we have here, we have to rely on the demographic information to tell us which people of Hispanic -- with that ethnicity are responding to it. But we don't have a designation as to whether they were given a Spanish-language version or not.

But as Joan said, the bottom line here is is that the uptake use on this has not been huge and it has been relatively isolated. There are single hospices that have a larger enough Hispanic population that speak Spanish or are more comfortable with Spanish and that there is some utilization. But we did not get

enough response back that we were able to analyze those separately even though we requested that they identify and let us know what they were doing.

CO-CHAIR PARTRIDGE: Sherri and then Lisa. Sherri?

MEMBER LOEB: Just a quick comment. I don't think the fact that it's not available in Spanish or any language should be a reason to not allow it to go forward, because if there is a population that's not English-speaking that can benefit by it, there are language lines. There are ways to eventually move it to another language. But in the meantime, by not allowing it to move forward you're hurting the larger population that can use it now. So, I mean, in the future we can move it forward and we can use it to utilize in the population that doesn't speak English by other means.

MEMBER MORRISE: So if I'm correct in thinking we're looking at its validity to

the reported population, not necessarily the validity to an expanded population that we would like to see measured in the future.

DR. PACE: That's correct. I think the question that comes up is is that clear, and that may be something for discussion in terms of if a hospice is taking care of a primarily Spanish-speaking population or non-English-speaking population, are they just not going to be included, or how do you interpret that? But I think you're right, we need to focus on the survey as how it's being used and the reliability and validity that they presented.

One other thing about the validity is that we also consider threats to validity here in your overall vote, which means case mix adjustment, exclusions, whether it can differentiate, whether there are different modes or different languages. So it's a lot that goes into validity.

CO-CHAIR PARTRIDGE: Chris?

MEMBER STILLE: Yes, I just wanted to quickly echo what Sherri said. I think measure developers can get excited when things get endorsed and that can actually spur them on to testing it in other populations.

I had a question also. When NQF endorses a measure are there sort of things that go along with it saying we endorse this, and by the way, it's probably only valid in XYZ population?

DR. PACE: That's a good question. We don't really explicitly state that, though we are endorsing the measure as specified. And I think part of the question is should this be more explicit that it's English language in terms of its primary use so that people better understand that? And it's a good question. It's probably something we need to consider for these patient-reported outcomes more carefully so that we can have some explicit information about that.

CO-CHAIR PARTRIDGE: Becky?

MEMBER BRADLEY: Yes. Well, I do think it's a good question and I think it doesn't just apply to the patient-reported outcomes, because we are seeing where NQF-endorsed measures for one setting are being applied to another setting. And so, I do think that's something NQF -- and that's probably a discussion for a different time, but I do think that is very important, that NQF consider the population that it is applicable for and look at how those measures are then expanded to other settings where they've not been validated.

MEMBER MORRISE: Well, I just wanted --

CO-CHAIR PARTRIDGE: Sherri, is your card still up?

MEMBER MORRISE: Yes, I just wanted to do a quick follow-up. So we're voting on the validity to the population specified, but the population isn't necessarily specified in explaining this to

the public.

The second part of my point would be that while providing this exact same survey in a Spanish language version wasn't discussed in the validity and reliability in terms of looking at how that may vary from the overall population wasn't explained -- that would have been nice, but my suspicion is that a Spanish language population has such significant cultural differences in terms of death and dying that a measure should be made specifically to address their population and cultural background.

CO-CHAIR PARTRIDGE: Len?

MEMBER PARISI: I had a blend of what Lisa was saying. I think cultural norms is such an important factor in hospice and end of life care that it is important to recognize that when an instrument is only in English, it does only address that population. I mean, there can be inferences, but cultural norms are so important in end of life care.

And practically speaking, there are two representatives here from large hospices in New York City and English works for us, but also we need Chinese, we need Russian and we need Spanish. So I think you can't have this discussion without acknowledging the fact that it does not reflect the entire population that we serve, and this is an important issue for that population as well.

I'm not saying that it's not something we should be endorsing because it doesn't reflect that, but I think it needs to be acknowledged that it's an English-speaking instrument.

MEMBER DOWDING: Sorry, can I just add one more comment to that? I think it's not just English-speaking. There's also issues of culture within that. And the point I was making is 93 percent of the respondents identified themselves as white. So it's not even that it's cultural in terms of language.

It's also culture in terms of other cultures and ethnicities. So that was my concern, is that it's valid for a certain type of population.

CO-CHAIR PARTRIDGE: Carol?

DR. SPENCE: I'm sorry, was there a question there? I heard that as a comment. Could you repeat, please?

CO-CHAIR PARTRIDGE: No, this is for one of our Committee Members, Carol.

MEMBER LEVINE: Yes, I just wanted to say we at the United Hospital Fund have a series of family caregiver guides including one on hospice and palliative care, and they're in four languages. And we have found that the Spanish language version, which is a generic Spanish language version -- it doesn't take into account the people of Spanish origin in California -- it's generic -- is as often downloaded as the one that's in English. So I think we can take this to an extreme of every specific cultural nuance.

The main thing is does it work for most of the people in that -- whose primary language is something else? Because the cultural norms may vary from one Hispanic group to another Hispanic group, certainly. And the Russian population definitely does that. Chinese as well. So I think it's important to recognize, but not to obsess about it. And I think that it is mostly important to acknowledge that the testing was done mostly on English-speaking.

CO-CHAIR PARTRIDGE: I'm looking at the clock and I think we all need a break. So are we ready to vote, or do you think -- I'm looking to staff for judgment here. Can we complete this do think in the next five to seven?

(No audible response)

CO-CHAIR PARTRIDGE: Yes? Okay.

DR. PACE: So I think these are good points that are brought up about the validity. It's definitely population-

specific. And I think the question is in terms of validating an instrument, it does make a difference that you include the populations that are going to be using it and sometimes it requires -- if there are consistent differences in how they respond to items that may require some adjustment of the score based on patients with certain language or cultural differences.

So I think in terms of your voting on validity, it should be based on the measure as its specified. And for the target population it may be that's something that needs be more specific in the target population, that it's for English-speaking population, if that would address some of the concerns. And I'll just ask those who raised that whether that would be of interest.

(No audible response)

DR. PACE: So Lee's saying let's go ahead and vote on it as its been specified and then if there are issues, we can come back

to that. I think that's probably the best approach. So on validity?

MS. ALLEN: So we're voting on validity. One, high; two, moderate; three, low; and four, insufficient. Starting the vote now.

(Voting)

MS. ALLEN: And the results are in. Seven high, nine moderate, one low and zero insufficient.

DR. PACE: Okay. So why don't we move on to feasibility and then usability? So basically for the surveys, these are not going to be in the EHR and generally require additional effort to submit them, but we'll ask the group that kind of looked at feasibility or anyone who wants to make a comment about feasibility for discussion, and then we can vote.

CO-CHAIR PARTRIDGE: Debra?

MEMBER SALIBA: Can somebody tell me what the response -- this is a mailed

survey -- what the response rates are? Do you get that data back from the hospice agencies that mailed them out?

CO-CHAIR PARTRIDGE: Carol, did you hear that question?

DR. SPENCE: Yes. So our response rate runs around 40 percent. It was a little higher back a few years ago. And as Matt said, we've had other -- we're curtailing of new hospices join. And then we made some changes to the survey a couple years ago and we saw a dip in the response rate. And it's come back up a bit. But it's around 40 percent, plus or minus, on any given quarter.

CO-CHAIR PARTRIDGE: So, Carol, is that 40 percent of hospices or 40 percent of people to whom hospice --

DR. SPENCE: It's the return rate --

CO-CHAIR PARTRIDGE: Okay. Thank you.

DR. SPENCE: -- on the mailed

surveys.

CO-CHAIR PARTRIDGE: Thank you. Any further comments, questions?

(No audible response)

CO-CHAIR PARTRIDGE: Ready to vote on feasibility?

MS. ALLEN: So we're ready to vote on feasibility. One, high; two, moderate; three, low; and four insufficient. Starting the clock.

(Voting)

MS. ALLEN: Thank you. And the results are in. Nine high, seven moderate, one low and zero insufficient.

DR. PACE: Okay. Then we move on to usability and use. And here we want to know whether it's in use especially for an accountability application based on what we saw in the submission form that it's not used for public reporting. It's in use, but my understanding is that the public could not look up a particular hospice to get their

score on this.

And, Carol Spence, that's correct?

DR. SPENCE: Yes, that is correct.

DR. PACE: Okay. Comments, questions about --

MEMBER VAN ZYL: I have a question for the developer. Is it possible for a family to call the hospice and request that information if they can't get it online?

DR. SPENCE: About a particular hospice?

MEMBER VAN ZYL: Yes.

DR. SPENCE: Is that what you mean? Well, I mean, some hospices put some of their scores out in their information that they have online, but that is up the hospices to determine whether they want to put this out or not.

The thing that NHPCO -- our approach philosophy on public reporting has been that public reporting of the survey really needs to come from CMS. We cannot

enforce, because this is not -- CMS for the CAHPS survey is requiring third party or vendor survey administration. We cannot require that. And so, a lot of hospices do their own mailing. A lot of hospices do their own submissions. And so we cannot enforce with the same authority as CMS accountability. So to have the playing field completely fair we decided not to go the public reporting route with this.

On the other hand, some of the vendors that have chosen to do this did some of their own ranking of hospices and gave that out to the hospices to use for marketing, and we did not approve of that. So this has been a quality improvement measure. This is the way we presented it. And again, CMS gave very serious consideration to adopting this measure when the quality reporting program was developed and decided to go with Hospice CAHPS. So we wanted to continue endorsement, as I said earlier in the introduction, for

this measure up until the time that Hospice CAHPS has gone through the endorsement process and we wanted to make sure that hospices that were using endorsed measure if they used FEHC.

CO-CHAIR PARTRIDGE: Thank you, Carol.

Len?

MEMBER PARISI: This is for the developer. I know that a lot of hospices report their data to NHPCO. Do you provide any benchmarking or reporting to the public if --

(Simultaneous speaking)

DR. SPENCE: Absolutely. We have very in-depth quarterly reports for them, and then we also do a national-level report that's an annual roll-up.

DR. PACE: Okay. So just a note about usability and use; and actually I should have said this about feasibility, it's not a must-pass criterion. And so, in terms of being used for public accountability and

improvement, you can vote it high, moderate or low in terms of -- it's not going to necessarily be a show stopper. And at the end we'll have one vote on your final recommendation. But I just wanted to let you know that a low vote here does not mean that it will not pass. Same way of feasibility.

CO-CHAIR PARTRIDGE: Ester, do I see a card up?

MEMBER NEUWIRTH: So the improvement on this measure since it was developed is pretty insignificant it looks like. I'm just wondering. So it looks like there's an increase in the number of people who are using the measure, but there's no sign of improvement based on measuring. So I'm curious just about the value of the use.

CO-CHAIR PARTRIDGE: I think that goes back to Debra's question earlier where -- response of the developer -- and I'm sorry I'm paraphrasing for you here, but I'm kind of moving us along -- was that for those hospices

that had been using this for a while, they've been able to stratify their data and the trend is that they're improving. But the overall average changes is being held a little lower by the fact they have new ones coming in all the time. I hope I'm saying that accurately. Yes? Okay.

MR. HASKINS: That's pretty good, yes.

CO-CHAIR PARTRIDGE: Are we ready to vote?

(No audible response)

MS. ALLEN: We're starting the vote on usability and use. One, high; two, moderate; three, low; and four, insufficient information.

(Voting)

MS. ALLEN: All votes are in. Five high, seven moderate, five low and zero insufficient information.

DR. PACE: Okay. So we'll move onto the final vote, which is overall. Now

considering all of the criteria, do you recommend this measure as meeting suitable for endorsement? This is the Committee's recommendation and what would be put out for public comment.

MS. ALLEN: Overall suitability for endorsement of Measure 0208, Family Evaluation of Hospice Care. Starting to vote. One, yes. Two, no.

(Voting)

MS. ALLEN: All votes are in. Sixteen yes, one no.

DR. PACE: And I'll just mention, because there was a note on that slide, that this will be pending our final discussion about competing and related measures. So this means that the measure looked at -- individually you would recommend, but we will have to come back to it because we have two other measures in this space. So for now this is something we'll come back to.

CO-CHAIR PARTRIDGE: And now

playing dragon lady, we're going to shorten our break, if you all will forgive us, to 11:30. Then I assume we'll also delay lunch a little bit. We've got two measures in the next bucket. I think some of what we already talked about will come up again in connection with the care measure and probably we'll move along a little more rapidly as a result.

And I know the developers for those two measures are standing by, so I don't want them sitting out on the phone too much longer.

DR. TENO: Actually, this is Joan Teno, I have a really hard stop at five after 12:00 that I can't get out of because I've been out of town and I have --

CO-CHAIR PARTRIDGE: Well, that gives us 30 minutes to consider it, Joan, so we will do our best.

DR. TENO: All right. I'm booked literally all day.

CO-CHAIR PARTRIDGE: That's fine.

DR. TENO: Thank you.

(Whereupon, the above-entitled matter went off the record at 11:21 a.m. and resumed at 11:30 a.m.)

CO-CHAIR PARTRIDGE: We're going to move on to Measure No. 1632, Consumer Assessments and Reports of End of Life.

Katherine and Lisa and Debra, if we could have everybody sitting down, please?

Okay. We're going to start on 1632, the CARE, Consumer Assessments and Reports of End of Life. And we'll start with Joan Teno. If you'd give a couple-minute introduction to 1632?

DR. TENO: Sure. And I'm going to be very brief. So I was requested by Sean Morrison and a group of people to submit this measure to fill a void for having some type of measurement to look at a systems perspective on end of life care. So they wanted a measure that would allow you to examine the quality of end of life care irregardless of the setting

of care. So they wanted a measure that could go across settings.

This measure is largely been based on the research they have done using a mortality fall-back survey to study bereaved family members' perceptions of the quality of care based on the site of care and focusing on the last place of care in the last week of life.

I've been in discussions with NQF staff because I do have concerns in that in the past three years my focus has been on creating a new version of a FEHC and working with the rating team on development of the Hospice CAHPS survey that we are only now starting to engage with a large ACO, a large corporation that has 22 ACOs to now chart time to get this measure to be used in the ACO context. So I think there's a concern in that we haven't done the reliability testing at the unit that you would want.

DR. PACE: So, and let me just

mention to the Committee, as Joan mentioned, we had some conversations with them about that. And as I've told you and we've talked about several times, the new guidance is to have testing at both levels. But because this is the first time we're applying those new criteria, we asked Joan to let the Committee evaluate the measure and then we can see where we are after you discuss the criteria whether it will stay in the mix or whether they will withdraw at this point given that they're working on another project. But we wanted to really provide the opportunity for the Committee to weigh in on applying those new criteria and whether you had any concerns about it.

CO-CHAIR MERLINO: So any members of the Committee that reviewed this have any general comments about the measure?

(No audible response)

CO-CHAIR MERLINO: The developer?

DR. PACE: No, that's okay. Why

don't we go ahead and move onto the evidence, which is again the same question: Are there health care structures, processes and outcomes that influence this experience? And there I guess was a little mix-up with the submission and there's a -- Joan sent a separate table; Nadine, if you could bring that up, where they did a crosswalk of preferred hospice practices that related to this measure. But we'll bring that up. And it's also available for the Committee on SharePoint, but just wanted to show you this. And we'll ask, like Jim said, the group that reviewed this: Carol, Brian, Sherri, Jim, Carin in terms of health care practices that influence this, if there's anything you want to say.

CO-CHAIR MERLINO: Peter?

MEMBER THOMAS: I saw this phrase a number of times throughout the different measures we looked at, that the Committee will need to substitute its judgment that at least one health care structure process,

intervention or service influence the experience of preparation for self-care.

So I just wondered; I know that each measure needs to relate to that somehow, why aren't the requesters making that clear and what does that exactly mean for us?

DR. PACE: Right. So in this case they did end up submitting this, so they do have what they considered as appropriate. And so, that was an initial comment, but they did submit the additional information. But what it means, the reason we put that in there is when they hadn't, I guess the thinking is that this Committee could make a judgment on that for these PRO measures, whether it is something that the health care unit can affect.

In terms of why the developers didn't submit it, I think perhaps they didn't understand the question. And we had had several conference calls and tried to provide some guidance in advance, but in some cases

there was probably a misunderstanding of what we were asking for.

CO-CHAIR MERLINO: Sherri?

MEMBER LOEB: In reading this and from the group call that we had it seemed -- and they had even said so, that this was more of a QI project rather than outcomes or a -- that we were going to be voting on. So that was my question. And I don't know if that comes here.

DR. PACE: Probably not here, but it is certainly a consideration when we get into usability.

MEMBER LOEB: Okay.

DR. PACE: And that actually is going to come up in the reliability --

CO-CHAIR MERLINO: Estee, do you have a comment? Your card is up.

MEMBER NEUWIRTH: (No audible response)

CO-CHAIR MERLINO: No? Lee, do you have something?

CO-CHAIR PARTRIDGE: Sherri actually raised my first issue, and it's not -- we'll get there.

CO-CHAIR MERLINO: Any other comments about evidence? Deb? Sorry.

MEMBER SALIBA: I think this table does a really good job of tying together sort of the rationale behind the structure-process-outcome links for each one of these. So thanks to the developers for sending it.

CO-CHAIR MERLINO: Great. Any other comments, evidence?

(No audible response)

CO-CHAIR MERLINO: All right. Let's vote.

MS. ALLEN: So we're going to start the voting on Measure 1632, CARE, and we're starting with evidence. One, yes. Two, no.

MS. DORIAN: And, Liz, you can send your vote via chat.

MS. ALLEN: Starting the clock.

(Voting)

MEMBER MORT: I just sent it.

MS. DORIAN: Great. Thanks. Got it.

MS. ALLEN: Please vote again. We're still missing --

(Voting)

CO-CHAIR MERLINO: Seventeen plus this one, right?

MS. ALLEN: Okay. Please vote again. Sorry. Please point in my direction.

(Voting)

MS. ALLEN: Okay. Eighteen yeses.

DR. PACE: Okay. So we'll move on to performance gap. Nadine, you want to bring up this -- okay.

CO-CHAIR MERLINO: Debra?

MEMBER SALIBA: So for performance gap I think there's clearly a performance gap in this area, particularly because we're not looking just at hospice agencies, but we're looking across different care settings. And

there is some information in here also about -- I don't know -- I'm a little confused about the divisions where it comes up about the cultural and ethnic gaps in performance. Does that come up here, or someplace else?

DR. PACE: Yes.

MEMBER SALIBA: Okay. So there's also some evidence of differences by ethnicity in terms of performance on this measure. And I think that's a really important point to bring out.

CO-CHAIR MERLINO: Any other comments, performance gap?

(No audible response)

CO-CHAIR MERLINO: Very similar to issues raised in the last instrument. Any other points?

(No audible response)

CO-CHAIR MERLINO: We ready to vote?

MS. ALLEN: Okay. We're starting the vote for performance gap. One, high; two,

moderate; three, low; and four, insufficient. Starting the clock now.

(Voting)

MS. ALLEN: The results are in. Fourteen high, four moderate, zero low and zero insufficient.

DR. PACE: Okay. So priority. This addresses whether it's a high-impact aspect of health care, affects large numbers, consequences of poor quality, and also in 1C5, this is specifically for PRO-PMs, how patients or families were involved in identifying that this is something of value and meaningful to them.

CO-CHAIR MERLINO: So this issue was specifically raised on the call. I don't remember who did question it. Does anybody want to talk about it or have any questions about it? Any comments?

DR. PACE: And the submission form I think said they conducted extensive focus group testing, but no specific information.

But again, that's the area we'll talk more about later. But does anyone want to make any comments or move on to vote?

CO-CHAIR MERLINO: Questions?

(No audible response)

CO-CHAIR MERLINO: Okay. We'll vote.

DR. PACE: So the question is --

(Simultaneous speaking)

DR. TENO: Can I just say that there is a published paper that reports all our focus group testing, and if you want me to summarize that, I can summarize it. I probably just didn't do it because it was just a very bad spring for me.

CO-CHAIR MERLINO: Does anybody need that? I think we're okay.

Debra, were you going to make a comment?

MEMBER SALIBA: I was just going to say, so we're talking about whether this is a high priority area and I was going to point

to the paper that was in here, so that's fine.

CO-CHAIR MERLINO: Okay. Any other thoughts?

(No audible response)

CO-CHAIR MERLINO: Then we can vote.

MS. ALLEN: So now we're voting on high priority. One, high; two, moderate; three, low; and four, insufficient. Starting the vote now.

(Voting)

MS. ALLEN: All votes are in. Sixteen high, two moderate, zero low, zero insufficient.

DR. PACE: Okay. So we'll move on to reliability, and this includes precise specifications, so if there is anything unclear about the specifications to specifically note that. One of the things that we asked that didn't get with the initial submission was the actual survey, so that is in your materials on SharePoint. And then of

course reliability testing which was done at the patient level instrument. And as Joan mentioned, they haven't had time to address performance measure testing because of not having that data from implementation.

So again, this is one measure. It's an overall composite as specified. And based on our current guidance about having testing at both levels, we would essentially say that this would be insufficient information, but we'd like to have some discussion with all of you in terms of whether you have any concerns about our current guidance.

As we mentioned, we're intending to endorse performance measures that are used in accountability applications, and that was the thinking behind the PRO project group in terms of why these should have testing at the computed performance score level. So I'll just stop and see what questions or comments people have.

CO-CHAIR MERLINO: Dawn?

MEMBER DOWDING: Can I just ask the developer for clarification? Is it right that the data on reliability and validity was actually collected in 2001 and '2 and it hasn't been updated since then?

DR. TENO: Yes, that's true.

CO-CHAIR MERLINO: Anybody else?

(No audible response)

CO-CHAIR MERLINO: Karen, can you go over just why the group, the staff recommended that it was insufficient to meet the reliability algorithm?

DR. PACE: Right, so the algorithm doesn't specifically address the PRO-PMs in terms of the fact that we're requiring testing at both levels. So it's a good point that the algorithm would say testing at one level would be moderate and testing at -- testing at the patient data level would be moderate and testing at the performance score would be high.

So the distinction is that the guidance for the PRO-PMs is that they must be tested at both levels. And the reason I'm suggesting that would be insufficient is we just don't know. We don't have the data to know whether it's reliable at the performance score level. We only have data at the patient level. So this something we'll work through together.

So the options would be to rate it insufficient because it doesn't have testing at that level, or to rate it moderate and we can note that that was only at the patient level and come back to that. I don't know if anyone has any other options that they want to bring up.

MEMBER BIERNER: So I just have a comment that the nature of this survey or this instrument is a variety of settings in which the patient lived at the last days of live, so it would be hard to know at what level you would test this at an institutional or agency

level. It seems like it wouldn't be clear to me which group you would go to for that because it's not just -- it could be in the home or in a variety of settings.

DR. PACE: And, Joan, can you remind us what level this was specified for for level of analysis? Is it any organization or was it specifically ACOs?

DR. TENO: So, the thinking was as we're moving to capitated payments that you needed to be able to test across setting of care. So ideally as we move forward with this, this would be an instrument that an NMA plan could use if they want to look at how they're caring for dying patients. And it also could be an instrument that an ACO could use. But I'm very honest with you that we're not totally there yet.

CO-CHAIR MERLINO: Any other comments?

DR. BURSTIN: In some ways because the guidance is so new we also -- we're a

little hesitant to push that it has to be met since it's literally the first time out of the box. So we want to at least present that information to the Committee saying the PRO Committee had just come up with this guidance fairly recently. This would be the very first time we have implemented it, so if the Committee wished us to allow them to move forward with a moderate, understanding additional work might be done in the future, that's something that could be done as well.

CO-CHAIR MERLINO: Peter?

MEMBER THOMAS: Is there anything the developer would like to say to address this potential concern about an insufficient rating on this item?

DR. TENO: The only thing I can do is be very honest with you where we are. We only have testing at the individual level. I think we have sufficient evidence at the individual level. In the future we need to address this concern. Can I give you any

evidence beyond what I have right now? No.

CO-CHAIR PARTRIDGE: This is Lee. I wanted to ask Karen and Helen a question here. This measure is similar to the one we discussed earlier this morning, and we know down the road not too far we're going to be getting a CAHPS Hospice survey. Would it be likely when we review the CAHPS Hospice survey later on that this measure also would come back for consideration at that same time?

So if we said go ahead and do a little more work on this and let it go through, it's almost like doing a time-limited endorsement, but I didn't use that phrase.

DR. TENO: Can I address that a little bit, because any further work on this measure will be done within the context of working with the exact same team, and I was a member of the team, who created the CAHPS Hospice instrument. So from the onset I understand the importance of harmonization. You should realize the CAHPS Hospice tool,

which we're not discussing, but -- really deals with hospice. That this instrument must deal with multiple settings of care, so it's going to have additional content.

CO-CHAIR PARTRIDGE: Yes, Joan, I think we appreciate that. It's just a question of whether they, the NQF will group them together so that we can in essence take a second look at this measure at that time as well.

DR. TENO: Yes, and believe me -- I think what I tried to do and slapped in an email message to NQF staff is, you know, I'm not invested in keeping this measure per se, but I'm invested in moving fields forwards. So once we have a better tool, my strategy in life is to try to make that a better tool, that harmonizes, available.

DR. PACE: But, Joan, correct me if I'm wrong, didn't you also say that you were working on a CAHPS version of this CARE, the --

DR. TENO: The exact plan is now to -- now that, at least in my hopefulness, that we are done with CAHPS Hospice, it's now to merge this work that I've done on this instrument with the CAHPS Hospice instrument to create a new instrument. And that's why I wasn't afraid to withdraw this instrument. I know that we need to -- now that we're far along with CAHPS Hospice, I need to now go back and take the work that I've done in this instrument and bring it up to the wonderful work that my colleagues at CAHPS, that are working on the CAHPS Hospice team, has done to harmonize it.

DR. PACE: So I think the question for the Committee then is -- and as you mentioned, Joan has been fully aware of kind of the new guidance on criteria, and one approach is for them to withdraw this measure and then bring the CAHPS cross-setting bereaved family measure in when it's ready. And it really depends on whether this

Committee agrees with applying the criteria and the guidance that for NQF endorsement the measure should have testing at both levels, or whether there's any concern about that.

CO-CHAIR MERLINO: Any other thoughts? I think we have to probably use the criteria.

(No audible response,)

CO-CHAIR MERLINO: No?

DR. PACE: So why don't we vote on reliability, and that will kind of solidify where you're at with it. If you agree with the testing at both levels, it would be insufficient. And if that's the case, then we'll stop there and we'll be talking with Joan about the next steps with their new measure.

CO-CHAIR MERLINO: Yes, our sidebar conversation was just that, the criteria holds the measures to a higher standard and provides the guidance for all of us. Do you have a comment?

MEMBER PARISI: I just have a question. Was it established that at the level of delivery that it's been tested, or just at the individual level? So for example, a nursing home versus an inpatient hospice unit, or a hospital for end of life care. I didn't hear that.

DR. TENO: It was tested at the individual patient level and then within each setting category, although I didn't present the data there.

DR. PACE: But each setting category it was lumped at the patients, not individual nursing homes or -- is that --

DR. TENO: No. No, right. Right. Only at the patient level in each.

CO-CHAIR MERLINO: Okay. Why don't we move forward with the vote?

MS. ALLEN: So we're voting on reliability, and it's including the specification and testing. One, high; two, moderate; three, low; and four, insufficient.

Starting the clock now.

All votes are in. Zero high, six moderate, two low, and ten insufficient.

DR. PACE: Okay. No. No, it's not. It's more than 60 percent in the low and insufficient. Yes. Right. So we would stop here then. And, Joan, we'll continue conversations with you offline on next steps and definitely everyone is interested in having the new instrument and the performance measures with the new instrument.

DR. TENO: Yes.

DR. PACE: And I thank you for bearing with us. I know we had this conversation, but we definitely wanted to have more Committee discussion about it as well. So appreciate it.

DR. TENO: Sure. Yes, and my only urgency is just that I've been traveling and I have meetings from -- my day is not going to end until very long. So thanks a lot. I'll sign off. Thank you very much.

DR. PACE: Thank you.

CO-CHAIR MERLINO: Okay. Moving to 1623?

MS. DORIAN: So we're moving to Measure No. 1623 now, which is the bereaved family survey from the PROMISE Center. So do we have Mary Ersek or Hien Lu on the call to introduce the measure?

DR. ERSEK: Yes, we have Mary Ersek, Hien Lu and Dawn Gilbert. I'll introduce it. The bereaved family survey arose out of a research study that was funded by the VA to Dave Casarett, who actually also is probably familiar to all of you because he's also been involved with the FEHC and also worked with Joan.

But anyway, this was developed. The validity testing has been published, and we can talk about that further. And then in 2008, the VA made a huge investment of resources into the Comprehensive End of Life Care Initiative. And the bereaved family

survey at that point -- and it used to be called FATE, by the way, so Family Assessment of Treatment at End of Life -- but then it morphed into what -- when it became a quality improvement and a performance measure, it was the bereaved family survey. So the PROMISE Center was set up to administer the survey, collect the data and report out.

We've evolved over time, but -- and the goal of the PROMISE survey -- so I just want to put this in framework. Although it's a performance measure, it was always imbedded in a QI program. Our mission at the PROMISE Center -- and we oversee kind of everything BFS, is to reduce variability in the quality of end of life care, because our veterans are not likely to pick and choose what VA hospital they're going to go to. Our job is to ensure that they can get good care at every VA facility.

We also have a mission to identify and disseminate best practices, and that's why

we have always been integrally involved in identifying those quality of care indicators that are associated with higher bereaved family scores.

I think there's a little bit of confusion from the earlier phone call as to, okay, what in fact is our performance measure? And remember that just like the FEHC in other measures, the bereaved family survey was endorsed by NQF as a palliative care measure, and when it was, and at the time, it was a composite score. At some point in its evolution at the VA there was a discussion with the National Performance Measure Work Group who suggested that we go with the overall item, okay, alone, which is item No. 18.

And we did this for a couple of reasons. Number one, it was easier to interpret. In addition to the performance measure, we also report out the scores on the individual items. So our measure is like the

FEHC. We have a number of Likert scale items that identify more specific elements of end of life care. So we did that for interpretability.

We did it for consistency, because if we were to change any of the items -- because this is a living breathing tool. If it was a composite score, then we would be hard-pressed to compare -- let's say we added an item on burial benefits, which veterans are eligible for, and we didn't have that before, we would have a hard time looking at trends over time.

The third reason for doing it is because our overall items are very similar; admittedly not identical, to the overall item on the compare -- or on the care, rather, on Joan's instrument, as well as the FEHC. And we felt as though -- or it was felt by VA leadership that this was a good thing, that while it wasn't actually apples to apples, it was like McIntosh apples, to delicious apples,

as opposed to apples and oranges. I'll stop there. Any other questions about this instrument?

CO-CHAIR MERLINO: No, that's good. Any comments about evidence?

DR. PACE: So the measure is just the overall rating out of the whole survey, and at least the performance measure that they're submitting for NQF endorsement. And so, they did submit some additional information about structures and processes that influence the overall experience. So any comments about --

CO-CHAIR MERLINO: Questions?

CO-CHAIR MERLINO: All right. Shall we vote on evidence?

DR. PACE: So the question is is there health care structure, process, intervention or service that will influence the overall experience with care? Nadine?

MS. ALLEN: So now we're voting on evidence. One, yes. Two, no. Starting the

vote now.

MEMBER SALIBA: While we're voting, just sort of a point that I'd like to ask. Is there any way to change the title so that it indicates that this is VA? I think that this is going to generate confusion, that there are so many different measures out there, and if there's some way to help distinguish it, it might just help people.

DR. PACE: Yes, I think that's a good point. We actually have asked and suggested that these measure titles actually indicate that it's a measure and not a survey. And of course now this is the overall rating. So there might be some ways for all of the measures that we're looking at to be specific about that, but I think we can deal with that with the developer.

MS. ALLEN: The results are in for the bereaved family survey 1623. And it's 18 yes, and zero no.

DR. PACE: Okay. We'll go on to

performance gap.

CO-CHAIR MERLINO: Questions, comments about performance gap? There was additional information submitted, as well, for this criteria.

DR. ERSEK: I just wanted to, if I may, point out there was a question on the earlier measure. What has scores done? And I mean, part of it is we're just lucky in the VA to be tied into this huge network of data, but if we you look at our scores, the scores have consistently gone up, not by much, but again when you're looking at these kind of measures, HCAHPS or any kind of really global measure, even a nudge of one or two percent, we believe is significant.

The other kind of cool thing here is, although we certainly don't have a randomized control trial, we can link this in our national databases with -- as these scores have inched up, and these are national facility scores, as these scores have inched

up, we can also demonstrate across the board that we are providing more care in hospice and palliative care units, which we would expect to be better care. And this is throughout the VA, but we also look at facility level. We have increased the number of veterans who receive a palliative care consult in the last 90 days of life, etcetera.

CO-CHAIR MERLINO: I think this came up in the call as well, but it seems like there's enough evidence for the performance gap. Any questions or comments about it?

CO-CHAIR MERLINO: Is everybody okay with moving on? Len? Sorry.

MEMBER PARISI: Unrelated question. When we think about these measures and voting on them, by definition an NQF endorsed measure -- shouldn't we be thinking nationwide, rather than just one particular health system, as in this case?

CO-CHAIR MERLINO: That's a great question.

DR. PACE: You know, you should be thinking that it's a national standard, but there's VA facilities all across the nation. And I think the developer also submitted some information about there's more diversity in their population -- although albeit mostly male, that there is some diversity. So sometimes we've endorsed measures that are being used in public reporting in one state. So the key thing is that it could be used as a national standard, albeit right now it's being used particularly in the VA facilities.

But, Helen, I don't know if you want to add anything?

CO-CHAIR PARTRIDGE: But I can tell you that, as a matter of desire, it would be nice eventually to have one measure.

DR. PACE: And that will be a discussion when you -- for all the measures you recommend, if they're kind of in the same space, to have a discussion about whether we need multiple measures.

MEMBER LEVINE: There are specific questions that specifically relate to a veteran's experiences that might not be comparable in other populations. So I think that's valuable to be able to relate to how at the end of life veterans bring back their combat experience, or whatever it is. It's not similar.

CO-CHAIR MERLINO: Sure.

MEMBER LEVINE: So I think there's value in keeping that part.

CO-CHAIR MERLINO: Any other comments? Questions? Becky?

MEMBER BRADLEY: Just a question for my own education. Did the VA hospitals participate in the CAHPS programs? Are they required to publicly report it?

DR. ERSEK: Yes.

MEMBER BRADLEY: So if there is a CAHPS survey, they would be migrated to that as well?

DR. ERSEK: I don't represent

that, but we do use the CAHPS. And I believe there is a movement underfoot to put all VA hospitals on the national reporting.

MEMBER SALIBA: It's not required.

DR. ERSEK: But I would have to --

CO-CHAIR MERLINO: Yes, I don't think it's publicly reported.

MEMBER SALIBA: It's not required, but the VA, as the developer alluded to earlier, is trying to use measures that could be compared to non-VA facilities, and to use the nationally endorsed measures whenever possible. But it's not a requirement, so we can't assume that they would use it at this point.

CO-CHAIR MERLINO: Okay. Why don't we vote on performance gap?

MS. ALLEN: So we're voting on performance gap, data demonstrated considerable variation or overall less than optimal performance across providers and-or population groups. One, high; two, moderate;

three, low; and four, insufficient. Starting the votes now.

All votes are in. Ten high, eight moderate, zero low, and zero insufficient.

CO-CHAIR MERLINO: Okay. Moving on to high priority. Comments?

CO-CHAIR MERLINO: Anybody want to discuss it from the call?

DR. PACE: So they did indicate that they did qualitative interviews with family respondents in four VA facilities, so they involved the family members. Any additional comments?

CO-CHAIR MERLINO: Becky, are you -- oh, that's okay. Any questions?

DR. ERSEK: If I can just say, I mean, if you look, we responded to that specifically. We developed it very similarly. I'm afraid none of us is very creative. We also, of course, referred to the consensus guidelines, etcetera, in the development, which I think we have in our original

application.

CO-CHAIR MERLINO: Okay. Shall we vote on high priority?

MS. ALLEN: We're voting on high priority. One, high; two, moderate; three, low; and four, insufficient. And it addresses a specific national health goal, or priority, or data demonstrated a high impact aspect of health care for PRO target population values and finds meaningful. Starting the votes now.

All votes are in. Sixteen high, two moderate, zero low, zero insufficient.

CO-CHAIR MERLINO: Okay. Now we're going move on to reliability testing.

DR. PACE: This will also include if there are any questions about the specifications, if that's unclear.

DR. ERSEK: Do you want us to respond to the information we got on Friday afternoon about --

DR. PACE: About the tests? Maybe -- yes, why don't we wait and let the

Committee speak and then --

DR. ERSEK: Okay.

DR. PACE: Well, let me just ask you: we still did not have any testing at the performance score level, so the question is do you have that?

DR. ERSEK: I thought that's what we submitted.

DR. PACE: Okay. Well, then let's wait until we have the Committee talk, and then we can come back to you. Thanks.

DR. ERSEK: Okay.

MEMBER BEVANS: My question for the instrument developer is the denominator statement. I'm wondering if you could justify some of these exclusions, because it really does reduce the overall number of people who are eligible to be included in a performance measure, in particular death by suicide, or accidents, or deaths that occur in the emergency department, or any of the others.

DR. ERSEK: Yes, the key point

there -- and actually I'll be honest with you, it's not so much the suicide. The key thing is we're asking families to tell us -- to evaluate the quality of care that the veteran received in the last 31 days of life.

So while we don't require that someone's been in the hospital for 31 days, and certainly the system has cared for this veteran, we felt as though someone coming in who spent less than 24 hours in a VA facility in the last 31 days of life -- how much of that could be influenced, or would there be enough care for that veteran's family to judge the care in the last 31 days of life?

MEMBER BEVANS: What if they're an outpatient --

(Simultaneous speaking.)

DR. ERSEK: If their contact with the VA was two hours in the emergency room.

MEMBER BEVANS: Any consideration to simply asking have they received care in the past 30 days? I'm concerned of course

that --

DR. ERSEK: Yes.

MEMBER BEVANS: -- by removing some of these populations that, for example, we may be limiting -- for example, the suicide limitation, that we may lose some important feedback about veterans who are challenged with some mental health challenges, as an example.

DR. ERSEK: Yes. And, well, many of these -- well, anyway, we do know there are psychiatric burdens, because we have those data available to us. You know, I think it's a good question, and actually we've just switched to a mail survey. I think it's a good question to raise. They are relatively -- they represent -- in the whole scheme of things, remember we have the -- well, our population, we essentially sample our population with those few exceptions, so we have data on about 82,000 veterans and we have bereaved family survey data on close to 36,000

of them. So I think they're relatively small numbers. Nonetheless, we can look at these.

The other thing that we're able to do, because again we have this administrative and clinical data available to us, we're able to compare those veterans, let's say, who committed suicide. We can actually -- we start with who died in the VA facility. And we could do analyses on those. And I think your point is well taken. And if there is a significant difference -- I think what really -- what it's stumbled is we're trying to identify actionable strategies, and if somebody's in the ER for two hours, do we fault the -- or how can we improve that care? ,Like can we get a palliative care consult in? Is that appropriate?

CO-CHAIR MERLINO: Okay. I don't want to cut you off, but I think we have -- Deb, did you want to --

DR. ERSEK: Okay.

CO-CHAIR MERLINO: You were going

to make a comment?

MEMBER SALIBA: I was just going to ask also if they were in -- I think I was leading into what you were about to say. If they were in hospice, outpatient hospice, would they be excluded of deaths within 24 hours admission without a prior hospitalization of at least 24 hours in the last 31 days of life? So if they were in outpatient hospice, and then ended up in the hospital for their last 24 hours, they would be excluded from this sample?

DR. ERSEK: Yes.

MEMBER SALIBA: Okay.

DR. ERSEK: Because we don't do -- now, if you're talking about community hospice, yes, because we don't -- we pay for community-based hospice. We don't --

(Simultaneous speaking.)

MEMBER SALIBA: Right, at our VA facility we do both community-based hospice and we also -- our palliative care team

follows people in the outpatient setting. So if those people hadn't been in the hospital in the last 31 days of life, then they would be excluded from this sample?

DR. ERSEK: Yes, I mean that said, we are actually -- up until now this instrument has always been used inpatient. Of course most vets don't die inpatient, and so we are extending the bereaved family survey to other programs. We have to do it slowly because of workload issues. For example, we're moving out into our home-based primary care programs. So right now they're excluded.

CO-CHAIR MERLINO: Karen --

DR. ERSEK: But in the future we're actually going to include them.

CO-CHAIR MERLINO: -- can you talk about the survey?

DR. PACE: Right. So, Jim is asking about the testing. So, basically, what we identified in the submission form is that you reported a range of Cronbach's alpha --

DR. ERSEK: Yes.

DR. PACE: -- but this just -- your submitting on the only the single item global rating for the performance measure.

DR. ERSEK: Right.

DR. PACE: So we don't have a reliability at the data element level, which we could talk about. But we also didn't see anything for the performance measure level.

DR. ERSEK: Okay.

DR. PACE: So the computed -- so each patient gives a global rating and according to your specification this is -- all the patient data is aggregated. I believe it's an average.

DR. ERSEK: Correct.

DR. PACE: Or top box. I don't remember. But the point is did you do any reliability testing for those computed facility scores?

DR. ERSEK: We did --

DR. PACE: Like a signal --

DR. ERSEK: Well, if you're talking about the Cronbach's alpha, you are correct. And I apologize for that. For us in the VA it's hard to disentangle the one item from the rest of the BFS, because we use it all even though the performance measure is technically only the single item.

We have run -- when you say reliability, it wasn't -- so test-retest, no, we've not done test-retest. And I guess I would ask what would be appropriate, and what would be gained from that? We did do splits -- where is it -- I'm sorry -- split half reliability testing. And I can send you these results, but basically we separated our sample of about 40,000 randomly. The means for both of those groups were 59 and 59, so right on. They were the same.

DR. PACE: And that was for the overall rating?

DR. ERSEK: Yes.

DR. PACE: Okay.

DR. ERSEK: Yes.

DR. PACE: So that --

DR. ERSEK: They're global items.

DR. PACE: Right. Okay. So that could be reliability of the data element level, split half.

DR. ERSEK: Yes.

DR. PACE: But the next question is then for the computed score do you have reliability testing, which would be how you actually compute the facility scores. So when you compute a facility score, is it the average or is it -- let's see. Okay. It's based on the optimal response. So it's the percentage of patients who --

DR. ERSEK: Yes.

DR. PACE: -- chose the optimal response?

DR. ERSEK: Well, their family. So we have an overall score --

DR. PACE: Right, right, right, right. Sorry.

DR. ERSEK: Yes, I know. On the overall score the choices are excellent, very good, good, fair, poor.

DR. PACE: Right.

DR. ERSEK: We dichotomize that into excellent, and everything else.

DR. PACE: Right.

DR. ERSEK: And so the report at the individual and the facility level is the percentage of people or families of that facility who reported the care as being excellent.

DR. PACE: Okay. So that's what we're asking, if you've done any reliability analysis such as signal-to-noise for that computed facility-level score, the percentage of patients who rated the care -- or families. I'm sorry. Percentage of families who rated their family member's care as high, or excellent.

DR. ERSEK: No, not to date.

DR. PACE: Okay. So, again, we

would be in the same situation that we were with the care instrument in terms of your rating, but there's one caveat here. I mean, you have all the data. It's probably something that you could do. So perhaps we could talk with you afterwards and we can think about how to -- yes. So why don't we go ahead -- if --

CO-CHAIR MERLINO: Any other questions or comments?

DR. PACE: How about Deb and the other reviewers for the reliability, if you have any thoughts? Len or Deb or Don, anything to add to that?

MEMBER SALIBA: The only additional thought is to try to be a little bit clearer that this is for veterans that die in the hospital at this point.

DR. ERSEK: Yes.

MEMBER SALIBA: It's not for all end of life vets.

CO-CHAIR MERLINO: Right.

DR. ERSEK: Correct.

MEMBER SALIBA: But that's --

DR. ERSEK: Currently.

MEMBER SALIBA: Yes. But otherwise no other comments.

MEMBER PARISI: The other comment that I would make is the same thing, that it's specific to the VA system, but also not the same as the FEHCs, whereas it just looks at end of life care, not specific to hospice, which I believe is a variation from what I'm reading.

CO-CHAIR MERLINO: Any other comments?

DR. PACE: I mean, this one is end of life care, right?

MEMBER PARISI: Yes, versus hospice care.

DR. ERSEK: And also the inpatient. I think it's easy to think about inpatients being acute care and intensive care. We can actually break down our findings

or our results. It also includes our hospice and palliative units, some of which have patients there for many months, because ours are run differently than Medicare --

CO-CHAIR MERLINO: Okay.

DR. ERSEK: -- inpatient units.

CO-CHAIR MERLINO: Any other comments?

DR. ERSEK: The other thing is TLC.

CO-CHAIR MERLINO: Any other comments from the Committee? Questions? Okay. Let's vote on reliability.

MS. ALLEN: So we're voting on reliability and it includes precise specification and testing. One, high; two, moderate; three, low, and four, insufficient. Starting the clock now.

All votes are in. Zero high, seven moderate, four low, and seven insufficient.

DR. PACE: So this would stop

here. And just so that we kind of do a check on where people are in terms of the moderate rating, did that mean that you all wanted -- those who voted moderate wanted to proceed, even though they hadn't done the testing at the performance score level? Okay. All right. Just wanted to --

(Simultaneous speaking.)

DR. ERSEK: Can I just -- so specifically the test you're looking for, because reliability is kind of broad, you were looking for noise-to-signal analysis?

CO-CHAIR MERLINO: We'll take that discussion offline, because we're running a little bit behind schedule.

DR. ERSEK: Okay.

DR. PACE: We'll talk to you afterwards, because I think you do have the data that you could do that, and provide that. So we'll definitely work with you.

CO-CHAIR MERLINO: Okay.

CO-CHAIR PARTRIDGE: So that

brings us to lunch time.

MS. DORIAN: First, before that we're just going to do a quick check-in with the operator to see if there are any --

CO-CHAIR PARTRIDGE: I'm sorry.

MS. DORIAN: -- public comments. So, operator, if you could please open the lines for public comment?

OPERATOR: And if there are any public comments, please press star one on your telephone keypad.

And there are no public comments at this time.

MS. DORIAN: Is there anybody in the room who would like to make a comment?

DR. JAMES: Thank you. Tom James. I'm the chair of the Health Plan Council for NQF, and a member of the public here.

First, I want to congratulate you on a lot of very lively discussion here. This is an important topic. I was on the Dual Eligibles Work Group. The hospice criteria

and measures we adopted in that were all procedural ones. I think it is important to be able to have lines of patient and family experience to supplement that. I know that we're going in the right direction. We want to get good measures, but I think that that is going to be key. And I'd love to see that go back to the Dual Eligibles Work Group. Thank you.

MS. DORIAN: Thank you. And before we break for lunch, I did also just want to do a quick poll to see who would be interested in attending dinner tonight at 6:00 p.m. Anybody? It's about two blocks from here. It's Mio. It's on Vermont. And I can put the directions up.

DR. PACE: Do you want them to sign up?

MS. DORIAN: No, no, no. I just wanted to get a head count, and then I'll put some information up on the screen at the end of the day so you can find it. Okay. All

right. We'll see how everybody's feeling.

DR. PACE: Well, just come up and let Lauralei know --

MS. DORIAN: You can let me know, yes.

DR. PACE: -- during the lunch break.

CO-CHAIR PARTRIDGE: I suspect that many of us did as I did, and didn't think there was going to be a group dinner and make alternative plans. But for the future it is kind of fun to get together, so we'll keep it in mind.

MS. DORIAN: Great.

MEMBER THOMAS: Quick comment about public comment. It strikes that before we take the final vote on a measure, it would probably be a lot more meaningful to get the public comment then, before voting. And then if someone wants to say something about that measure, the vote has already been taken, it's largely water under the bridge. So is the

purpose of the public comment to get the public's view on these measures before we finalize our vote, or is it just if something comes up that they'd like to bring to our attention, generically?

DR. BURSTIN: It's a good question, Peter. It's probably more the latter for this point for public comment, but public comment will be much more significant following the draft report that goes out for comment.

MEMBER THOMAS: Okay.

DR. BURSTIN: People can submit online comments at that point. And keep in mind they can submit comments on measures you recommended, also measures you didn't recommend or any measures in the gray zone. So you'll get, potentially, the opportunity.

Anything you may have in fact said no to today, you could get a groundswell that puts additional information forward, including from the developers that could change your

mind. So this is still pretty early in the consensus process. So you usually get pretty minimal comments at an in-person meeting, but much more to follow.

MEMBER BEVANS: I know we're all anxious to go to lunch, but we're making a recommendation for approval, but what happens next? How does the approval process actually happen and integrate the public comment?

DR. BURSTIN: Right, so what will happen is is when we get the public comments back, we'll set up a post-comment call with all of you to review them. The comments will be themed by staff. You'll have an opportunity to see if there's any measures you want to revisit prior to a final report going out for vote. It will then go to our Consensus Standards Approval Committee that reviews all the final measures recommended. Actually Lee sits on that Committee. So you'll have an in-house rep there. And then ultimately to the Board for ratification.

And since we know this group will be ongoing in other phases, you'll have lots more opportunity to revisit. Even some of the measures that perhaps just needed more work could potentially even come back in later cycles.

DR. PACE: Okay. Thank you. So we have a buffet set up. We'll try to get back to our seats at 1:00. You can eat back there or bring your lunch to your seat, but we'll reconvene at 1:00.

(Whereupon, the above-entitled matter went off the record at 11:30 a.m. and resumed at 1:00 p.m.)

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

(1:00 p.m.)

MS. DORIAN: Okay, everyone, we're about ready to get started again. I hope you enjoyed your lunch.

Our next slotted item was going to be competing measures and harmonization, but we actually will skip over that because two of the measures in that discussion did not move forward earlier on, so we gain a little bit of time that way.

So we're moving on to 0258, which is the CAHPS In-Center Hemodialysis Survey from CMS. And we have the developers. If you'd like to introduce yourself.

MS. CRAWLEY: Good afternoon. I'm Barbara Crawley and I'm here for 0258, the In-Center Hemodialysis CAHPS Survey.

MS. GOLDSTEIN: And I'm Liz Goldstein. I'm from CMS, and under my division at CMS is the In-Center Hemodialysis Survey, the Home Health Care CAHPS Survey and

the Hospital CAHPS Survey.

Well, also there was a lot of discussion about the Hospice CAHPS Survey this morning and so the development of that survey has been completed. So we will be submitting it next time there's an opportunity to submit.

MS. DORIAN: Great, thanks. And if one or the other of you would like to give a brief introduction to the measure.

MS. CRAWLEY: Like all the CAHPS surveys, the In-Center Hemodialysis Survey stresses standardization, beginning with the training of our survey vendors, letting all of the facilities know that this is going to take place from a protocol that is standard across all of the process. And this is our training manual. This is our manual. This is our protocol and every CAHPS survey has a protocol that must be followed.

Even with the data submission of the survey results, it's all done in a template that must be submitted to CMS in that

format.

As the first CAHPS survey to address the experience of a specific patient population, the results from the ICH CAHPS work will address accountability for dialysis centers, drive quality improvement and provide performance measures for quality measures. And we hopefully will start seeing public reporting in 2016.

Currently the data that we have is based on the pilot testing that was done back in 2005. There are 5,800 dialysis centers with nearly 370,000 patients with permanent, irreversible kidney failure.

What we have out on the CMS website right now as far as quality data has to do with the clinical information surrounding dialysis care. There is nothing about the provider-patient interaction or the communication that goes on within a dialysis facility to assist the patient about information that they will receive over the

course of time.

One of the questions that just came from a facility that they want to add to the current survey, which they're allowed to do, is how warmly were the patients greeted when they first arrive at the center. So there is some interest, even at the facility level, about communication.

We have talked a lot about the interaction between the patients and the staff and we know that Medicare does have regulations that require dialysis facilities to implement a process for beneficiaries to file grievances. But what we've learned is a lot of times patients do not file these grievances. They may not know about the avenues that they have, or a lot of them fear their airing of their problems.

So a lot of our questions do address problems that they may encounter in the survey about communicating those problems. And I'll stop there, so you can ask questions.

MS. DORIAN: Great. So just to remind everybody who the reviewers of this measure were, we have, for importance, Becca Bradley, Carol, Lisa, Lee and Peter. And then for scientific acceptability, Sam, Dawn, Sherri, Len and Deb.

MEMBER BIERNER: Can I ask something?

MS. DORIAN: Sure.

MEMBER BIERNER: Just want to ask the developers, the issue of this reluctance of the patients to give adverse information or negative information, how has that played out? How are you monitoring that? Is there some way you've established that you're getting reliable responses, or at least in the format of your outcome measure?

MS. CRAWLEY: Well, in the protocol, the survey may not be administered within the facility. It has to be done by a third-party vendor who will contact the patient directly, either by one of three

modes, mail only, mixed mode, which is mail followed by telephone, or telephone-only mode.

We do have a question on the survey that asks, did anyone assist you with this, filling out the questionnaire? And we will eliminate those questionnaires where it said someone at the facility assisted.

MEMBER BIERNER: If they have a physical handicap that they need assistance, that's not disallowed?

MS. CRAWLEY: That's not disallowed if it needs to be translated into a language that we have not standardized thus far. We have it in English, well, Spanish, Chinese and American Samoan right now, because there is a couple facilities in American Samoa.

CO-CHAIR PARTRIDGE: But I think, according to my notes from our work group call, am I not correct, this is a situation in which you got a special protocol and you pull the sample and you notify the patients, Liz?

MS. GOLDSTEIN: Yeah, so for this, for some of our surveys, the provider provides a discharge list or current patient list to the survey vendor.

For this survey, when we get to national implementation this fall, CMS is pulling the sample. So we want the provider totally separate from the survey process.

We also, for this survey, because we knew of some of the concerns, we did over the winter a lot of testing of our cover letter with ESRD patients to make sure we're giving them any reassurances they need in that cover letter.

CO-CHAIR PARTRIDGE: So I think that, as I heard that, one of the questions for this committee to think about is, is that protocol a specific part of the specifications for this measure? Because, as I recall, you were extremely concerned about whether patients would be comfortable responding.

MS. GOLDSTEIN: I think, for us,

yes. I think for the -- and I'll maybe make this point now, but I think it's the same for the ICH CAHPS Survey, Home Health Care CAHPS Survey and the Hospital CAHPS Survey we're talking about today.

We have, for all three surveys, very standardized protocols. And NQF had given us guidance. We don't need to go through all our specifications. We should refer to what we have online on our various websites.

So for all of our surveys, we would say our full protocol, which are in these thick manuals, is our specification. I think as we, CMS, has implemented surveys nationally, we have found that, you know, just giving measure specifications and questions are not sufficient to ensure standardization of a measurement.

So I would have great concern, you know, just implementing it from a survey without sampling protocols. When you do the

survey. What's a telephone complete? What's a call attempt?

Our things are very, very, very specific, much more specific than what we've been talking about earlier today. And so those really are part of our specifications.

DR. PACE: And NQF would consider, as you saw, that those were items in the specifications, sampling and survey instructions.

Because of the voluminous manuals, we didn't want them to copy and paste that into the submission form, and suggested that they could put, you know, kind of some key points in the submission form and refer to the manual.

So there's only so much that everyone can handle here. But we definitely consider sampling and survey part of specifications.

So maybe we can get on to 1A, the evidence. And, again, this is about are there

interventions, services, actions by the, in this case, the dialysis facility that will influence the experience with care?

And those of you who reviewed this part of the measure may want to make some comments. Microphone, please.

MEMBER BRADLEY: Since dialysis is being provided now in multiple settings, patients are coming in as outpatients, some are receiving dialysis as units in a hospital, is there any variation or is that risk adjusted in any way? Are you looking at the experience based on the setting of where the dialysis is occurring?

MS. CRAWLEY: Part of the protocol is that our inclusion criteria, it must take place in a in-center hemodialysis center as an outpatient.

If you're having acute condition and you're already on dialysis and you must continue dialysis while you're in the hospital, you will not be selected for the

survey. So it's only patients receiving in-center hemodialysis care as an outpatient.

MEMBER PARISI: That actually answers my question, because I was actually enthusiastic that this would be a first measure that would measure care across the continuum, but that's not the case.

But it is only reflective of a population with a diagnosis of ESRD, is that correct? And the sample will be selected, I'm assuming, based on the CMS criteria for billed visits for dialysis? Is that correct?

MS. CRAWLEY: Well, they have to have been at a outpatient center for three months or more to be selected. They have to have end-stage renal disease. They have to be 18 years or older.

If you're in a nursing home and you're receiving dialysis care where you're going outside of the nursing home to have that, we're excluding those patients because we feel as though they'll be confused about

their care, whether it's nursing home care versus the dialysis care.

MEMBER PARISI: So it's ambulatory?

MS. CRAWLEY: Absolutely.

DR. PACE: So any comments about the influence of the dialysis facility on the experience that's being measured?

MEMBER THOMAS: Correct me if I'm wrong, but the vast majority of ESRD treatment is performed in hemodialysis centers on an outpatient basis. It's something like over 90 percent or something, right?

MS. CRAWLEY: That's correct.

MEMBER THOMAS: Peritoneal dialysis is the other major method and that's relatively small.

MS. CRAWLEY: A small percentage.

MEMBER THOMAS: So we're getting to the vast bulk of those patients and measuring it.

MS. CRAWLEY: Absolutely.

MEMBER THOMAS: Is there a reason why, and did I miss this in your introduction, is there a reason why we need to substitute our judgment for the tie between whether this impacts a structure, process, intervention or service?

DR. PACE: Oh, that was my note based on what was actually provided in the submission form. So they provided some information about the types of content that was included in each of the performance measures and discussed some general healthcare actions.

And so the question to the Committee, because they didn't do kind of a one-to-one, is are you satisfied that the dialysis facility can influence these?

MS. GOLDSTEIN: You know, we could add something to that. For all the CAHPS surveys, the patient experience items are all -- I mean, that's part of the development of the survey -- are all things that the provider

has control over and they're all things that come from the patient, things that they express to us that they want information about.

So the items that we included on all these surveys are things where the patient is the best or only source of information. How the survey is developed, all the items are things that a provider can influence.

DR. PACE: And I think for these, you know, the questions that go with a performance measure are kind of indicative of the things that the provider can do. But, you know, it is just a question to the committee if you agree.

MEMBER LEVINE: Since the pilot, am I right, the pilot data was collected 2005?

MS. CRAWLEY: That's correct.

MEMBER LEVINE: And it's 2014. Is there a reason why it took so long? Or is there any reason to think that that pilot data might be outdated in any way?

MS. GOLDSTEIN: I think the survey was developed a number of years ago, and there's been a lot of discussions within CMS for a number of years about moving to national implementation. Some of it had to do with budget issues and, you know, different priorities.

But CMS, right now, we have a huge emphasis on developing and implementing patient experience of care surveys, so we have a number of additional ones we're currently developing.

So as part of this, you know, increased emphasis on patient experience, there was a decision to move forward with national implementation of this survey.

The data that we're submitting to NQF at this point is the same data we submitted in our original endorsement. We're not going to have new data until January 2015 with the first administration of the survey with reporting to CMS.

Over the last couple years, facilities have been required to use the survey, but to do it on their own, not using approved vendors, the protocols haven't been so rigorous and they did not have to submit the data to CMS.

MS. ALLEN: We're starting the vote on evidence. Rationale supports the relationship of the health outcome, or PRO, to at least one health care structure, process, intervention or service. One for yes, two for no. The timer starts now.

The vote is in. Eighteen yes, zero no.

DR. PACE: Okay, let's move on to performance gap, 1B. And this can be performance gap at the facility level or disparities.

MEMBER SALIBA: So I'm on the post-acute care/long-term care NQF panel and we have been looking at measures in the area of dialysis, and there is a huge, huge, huge

need for measures focused on this population. They're very vulnerable. They're at very high risk for -- they have multiple, as it says in the documentation that was provided, they have multiple comorbidities, often some cognitive impairment.

So this is a really important area where there are huge gaps across dialysis centers and providers in terms of the quality of care that's being provided.

MEMBER THOMAS: With respect to disparities, my understanding is that this population is very disproportionately minority population. Is that correct?

MS. CRAWLEY: Yes, that's true. When we think about the comorbidities that exist in this population, the diabetes, heart disease, those are conditions that you find predominantly in African American community.

MEMBER THOMAS: So --

MS. CRAWLEY: If you look -- I'm sorry. Go ahead.

MEMBER THOMAS: So the fact that, reading in some of the notes that have been prepared by staff that there is no data on disparities in the experience of care, but isn't just collection of the data itself reflective of a fairly significant portion of African Americans being sampled, correct?

MS. CRAWLEY: Right. I think in our results from the pilot testing we had about 40 percent African American, 40 percent white and the rest other.

So we did not sample by race, ethnicity, but we captured, say, 40, you know, probably an even split between African Americans and whites.

MEMBER THOMAS: Just by virtue of the population?

MS. CRAWLEY: Just by virtue of the population.

CO-CHAIR PARTRIDGE: I'm sorry, I should have asked. This is an all-payer, am I right? This is not just Medicare patients?

MS. CRAWLEY: It is from centers who are --

CO-CHAIR PARTRIDGE: Medicare-certified, but it's their entire caseload?

MS. CRAWLEY: Right, 18 years old and older.

CO-CHAIR PARTRIDGE: I'm a former state Medicaid director so I suddenly thought, oh, if you get current data, you're going to get a lot more Medicaid patients in there.

MEMBER THOMAS: Are you talking about the high percentage of Medicare beneficiaries who are ESRD?

CO-CHAIR MERLINO: Yeah, I was just saying that I think most end-stage renal disease is Medicare, right? That's one of the few --

CO-CHAIR PARTRIDGE: It makes you eligible for Medicare. There are people who are not Medicare eligible.

DR. PACE: So we don't have performance data on the measure as specified

in the section, but the question is, you know, obviously, as Deb Saliba was mentioning, this is an area that's thought to have areas for improvement, and I guess see if anyone has any other specific questions or comments to add to that before you move on to voting.

(Pause.)

MS. ALLEN: So we're voting on performance gap data demonstrated considerable variation or overall less than optimal performance across providers and/or population groups. One high, two moderate, three low and four insufficient. Starting the votes now.

All votes are in, ten high, seven moderate, one low and zero insufficient.

DR. PACE: One thing I want to just remind everybody about now as we continue on, and I should have said this at the beginning, this is not one measure and we're not endorsing the survey.

So there are six performance measures that you're actually looking at here.

So it's nephrologist communication and caring, the quality of dialysis center care and operations, providing information to patients. And then three global items: rating of the nephrologist, rating of the center staff and rating of the facility.

So just keep in mind as we're going through this and start getting specific information about those individual performance measures, that if there's any one in particular that you have a question about, you know, we can discuss that.

But if not, we'll vote on them kind of en bloc. But just wanted to call that to everyone's attention because we've kind of switched now into these submissions that have multiple measures.

CO-CHAIR PARTRIDGE: Karen, just to be clear, if you think that one of these measures is either very strong or very weak and you wanted to vote on it individually, we can pull it out and we can do that.

DR. PACE: Yes.

CO-CHAIR PARTRIDGE: Okay.

DR. PACE: Let's go on to high priority. I think we've had a little bit of this discussion already in terms of the patient population and some of their characteristics.

And the submission, I believe, talked about using focus groups and the developers here have also mentioned that. So comments or questions?

CO-CHAIR MERLINO: Anybody disagree that this isn't high priority? Should we just go to vote? Let's go to vote.

MS. ALLEN: So we're voting on high priority. One is high, two moderate, three low and four insufficient. Starting the vote now.

We're still missing a vote. Thank you.

All votes are in. Results show 17 high, one moderate, zero low, zero

insufficient.

DR. PACE: Okay, we'll move on to reliability, which includes the specifications and reliability testing at the scale or patient-level measure score and the computed performance score for the dialysis facility.

CO-CHAIR MERLINO: How many facilities in the United States do you estimate?

MS. CRAWLEY: 5,800 approximately.

MEMBER THOMAS: There were some red flags that kind of went off a little bit in this one, like in terms of reliability, like I guess the vulnerability of the patients, the statement a few times that many of these patients have significant cognitive limits, a number of them fear retribution in terms of sampling what they think of their own care, a number feel that they're lucky to be in a facility that has a time slot available that's maybe close to their house and if they complain maybe they wind up losing that and

wind up being sent across town.

And so how would you say that really impacts reliability of the data that was collected?

MS. GOLDSTEIN: I think how we've set up the protocols are to ensure that patients feel that they can answer the questions honestly.

Even from the field test, you know, a number of years ago, and Barbara and I worked on that field test many, many years ago, but there we felt like patients were honest. The scores, there was a lot of variation across, you know, facilities.

So people were not afraid to speak up, I think, as long as they're reassured that their responses aren't going straight back to the facility and the facility is going to act upon those responses.

So I think as long as the protocols are set up, they're rigorous and, you know, the patient feels comfortable and we

make them comfortable in responding, I think they will, you know, answer honestly.

CO-CHAIR MERLINO: Any other comments?

MEMBER DOWDING: I've just got a question. I'm going through the data on reliability and I can see the reliability for the three composite measures. But do you have any data on the three global measures? That seems to be missing in our submission, either that or I can't see it.

MS. GOLDSTEIN: No, it is missing. So Barbara and I were not the ones who submitted this original package years ago. So we also realized that that was missing from this submission.

So we had to track down who still has access to these data, because we do not have access to it. We have tracked that down and we will have the facility-level reliability for those three global ratings later this week.

So I think NQF said, to at least some of my staff last week, that if we have additional information we could submit it in August.

DR. PACE: Yeah. What we'll do is, you know, if there was additional information, I don't know if you have anything to report on that verbally, but there's a couple of things then.

One is, you know, whether that causes any concern in terms of moving this forward based on the information you do have, or you can go ahead and vote on this and they can submit it. During the comment period is usually the time that we can gather any additional information that's needed to supplement what was already submitted.

So, you know, what we identified here is the -- this is for the three scales, as Dawn was mentioning, and the ICHR is the facility level, the performance score level reliability. And then the last column, alpha,

is the scale level, the patient level scale, and all these statistics are good.

And I don't know if you have any general comments, or know anything about those others that you want to comment on or just --

MS. GOLDSTEIN: I think we have to wait til later this week. What we have access to is information, for example, the communication one, the overall rating. Right. That one about ten percent are rating kind of near that top box but there's a lot of variation. The other ones we know from the information that we have they're not topped out at all. Very few are weighing up high but we'll have more information later this week.

CO-CHAIR PARTRIDGE: Katherine. No, I'm sorry, Dawn.

MEMBER DOWDING: Yes, it's just a query. I don't know if I'm comfortable voting on three measures for which we don't have any reliability or validity data. It's just a personal thing. I'm quite comfortable voting

on the data, for the data we have, given that these are six measures, not one composite.

DR. PACE: That's fair, and we can deal with the other three when we have that information.

So if you want to comment on the reliability data that we do have for these, you'll see that, the In-Center, .64, .51 and .53, which are a little lower than the CAHPS usually likes. And, again, this is pilot data so I don't know if Barb or Liz want to make any comments about that or your thoughts about that.

MS. GOLDSTEIN: Yeah, I mean, one, this is pilot data so it's off of a small number of facilities, only 32. So these are the types of statistics that we regularly run on all of our data. So, you know, we'll re-run these again obviously when we get data in January and there'll be obviously a lot more data.

DR. PACE: So, you know, I'll just

put this in perspective. You know, so this is similar to what we would see for an initial endorsement, and for various reasons this was delayed getting implemented.

It's a reason that, you know, if this were the, you know, second time it came up we would like to see more robust information. But I think you should maybe kind of think of it in the perspective as you would an initial, you know, endorsement.

CO-CHAIR MERLINO: Peter.

MEMBER THOMAS: I have a question about one of the last things on some of the staff prepared notes. It says that this is due to be implemented for payment programs, meaning the ESRD Medicare payment program, in 2014 and '15.

So regardless of what happens to these measures in this process, is CMS going to be implementing this nonetheless? Or will this make it dependent upon whether they do that?

MS. GOLDSTEIN: CMS is moving forward with implementation. So there have been, in prior rules for prior years, there's been information about how facilities needed to do the survey on their own, not reporting to us. So that's already happened for two years.

And for 2014, facilities are required to collect this information. So it'll be the fall of this year. So we are moving forward, you know, with data collection. So it's not contingent on this endorsement.

MEMBER THOMAS: So the only distinction there then would be that CMS would move forward with a measurement tool that is not approved by the NQF. But they're still moving forward with it. Am I right?

MS. GOLDSTEIN: Correct.

MEMBER THOMAS: All right, thanks.

DR. PACE: Right. So generally CMS likes to utilize NQF-endorsed measures,

but they also have mechanisms to proceed when they need to.

But basically, you know, the Committee votes on the measures and the information that's presented, and then, you know, we try to sort things out as we go forward.

So I think, again, that's why I wanted to say, I think you need to think about this in the perspective of almost like an initial endorsement because it's still kind of the pilot testing data, and is it sufficient to move it forward for implementation?

And then, you know, in the next endorsement maintenance, have much wider implementation and data to say, well, you know, this hasn't improved. What's going on? Or see what the difference is as we have wider implementation.

CO-CHAIR PARTRIDGE: Carol.

MEMBER LEVINE: I'm still concerned about the question Peter raised

about the protocol and how it reassures the patients that they're actually -- the information is not going to be used against them.

And I'm wondering if there is any provision for staff training or who is -- if it's mail, I mean, some of these people are not going to be able to fill out something. Is it phone call? Is there training about how the facility tells the patient that you may be getting this questionnaire, this survey?

There's a lot of pressure. A lot of these, maybe most of these dialysis centers are for-profit centers and there's a lot of pressure to look really good. And there's just every incentive to make the results look good unless there's a counterbalance that says, you know, you have to reassure these patients. I may be making things up except that I hear all this stuff.

DR. PACE: Let's let the developer respond to that.

MEMBER LEVINE: Thanks.

MS. GOLDSTEIN: So the facility, besides contracting with a CMS-approved survey vendor, is not involved at all in survey administration procedures. So there's a set of approved survey vendors. How many do we have?

MS. CRAWLEY: Twenty-two.

MS. GOLDSTEIN: Twenty-two for this program, and these 22 vendors have been trained by CMS on all the protocols. These trained vendors will have 800 numbers, so if a patient gets a survey and doesn't know what to do, they will call these 800 numbers, which is going to a survey vendor and not to the facility.

So the facility could, and they're allowed to, let patients know they may be getting a survey, but that's it. That's the only interaction that the facility has with the patients over the survey results.

The facility will not get the

individual-level survey results back. They will get aggregate data. They can get it by patient characteristics as long as -- even by patient characteristics it's not enough to identify, you know, who a patient is. So cell sizes have to be, you know, 11 or more if we're doing, like, a cross tab by characteristics and responses.

So there's a lot of protocols that are set so the facility can never identify how, you know, Mrs. Jones responded on the survey.

MEMBER LEVINE: It would be good to have that kind of information in the submission. I don't think it was as completely as you've described it.

DR. PACE: So I'll just jump in here and explain that, you know, that's part of what we were talking about before. They have, you know, a very extensive protocol.

And this is something we can talk about tomorrow when we get to your feedback

and suggestions of, you know, if it's a 500-page manual, we can't have that copied and pasted into the submission form. And what's the best way, you know, what are the key things to communicate?

So, you know, definitely there's pros and cons and a myriad of ways and definitely think about that and if you have some suggestions so that we can talk about that tomorrow, because it's a challenge for these surveys that have very extensive protocols. Okay, so, Jim.

CO-CHAIR MERLINO: Any other questions? Sorry.

CO-CHAIR PARTRIDGE: Lisa.

MEMBER MORRISE: So I am just wondering if we, as a body, should be consistent in the application of the standards that have been set up from one to another.

And this particular effort, though certainly significant, does not have the data for the individual centers, right -- or the

centers. It's for the individuals only. Is that correct?

DR. PACE: No, what we have here is the patient-level data and the center-level reliability statistics for three of the measures that are based on multi-items.

So what's missing that we talked about, and the developers mentioned, is for the three global ratings they don't have it right now but they can get it to us.

So I think one of the suggestions was for the group to focus on the three measures for which there is data and then they'll come back to us with the other data. Does that make sense?

CO-CHAIR PARTRIDGE: Karen, just to clarify, I think, are we technically, from the perspective of voting, setting aside Measures 4, 5 and 6, which are the global measures, and either not voting on them or are we voting them insufficient today or low today or --

MEMBER MORRISE: So should we, in essence, as a body, vote on what we're voting on?

(Laughter.)

CO-CHAIR PARTRIDGE: Yes, well said.

DR. PACE: And maybe we can do this by a show of hands. So let me just test this out.

If you're in favor of kind of splitting the question so that we have you deal with the three measures for which you have all of the data separately from the three that we don't, if you're in agreement with that, would you just raise your hand to kind of split the question there?

(Show of hands.)

Okay, all right. So and maybe the cleanest thing to do would be let's vote on reliability based on the three that we have the data.

And, you know, I guess we could

formally vote on the others, which would be insufficient and would end for that right now. And then they could come back with that. Would that work for everybody?

So let's do reliability on the three that are here. And these three have testing at both levels. We've talked about, you know, the pilot, that maybe these aren't as high as you might see when it's fully implemented, you know, but it is tested at both levels.

So it would be eligible for a high rating at both levels, or if you think the statistics are maybe not as high as we would like, moderate rating is fine and either of those pass.

So I think -- we okay to vote? So you're voting on the three measures for which you have data on both levels.

MS. ALLEN: Reliability for three measures that have data on both levels. One high, two moderate, three low, four

insufficient. Voting starts now.

All votes are in. Five high, 13 moderate, zero low, zero insufficient.

DR. PACE: Okay, so now the question is how we vote on the -- why don't we just do a hand vote on that? Our software, we can't switch measures in the middle, so that's a little technical glitch but I think we can do a hand count on the three that we don't have the information. Go ahead.

CO-CHAIR PARTRIDGE: For the record, we're voting on M4, 5 and 6, rating the nephrologist, rating the staff and rating the facility.

DR. PACE: Yes.

MEMBER THOMAS: And a vote for insufficient would still allow the developers to bring additional evidence in the same cycle, as opposed to voting moderate and then they do the same thing?

DR. PACE: Right. Insufficient means that there's no data for you to rate it

against the criteria, which is where you're at.

But we will have a period of time during the public comment period where the developers could submit that information so that you could act on that after the comment period. Does that make sense? Okay. All right.

(Off microphone comment.)

CO-CHAIR PARTRIDGE: Well, that's actually kind of what we're doing.

DR. PACE: We can do that. All right, so maybe that'll be the -- why don't we just use that as the vote then? Perfect.

CO-CHAIR MERLINO: We're going to have the same problem with the other one.

DR. PACE: Well, this will stop here.

CO-CHAIR MERLINO: For all six?

DR. PACE: No, just for the three, just for three. Let's finish reliability and then we'll --

CO-CHAIR MERLINO: On to validity.

CO-CHAIR PARTRIDGE: Right. Okay, all in favor of deferring essentially. Okay. Got that? Do you want to repeat it just into the record?

DR. PACE: So, yeah, so actually everyone voted in favor of deferring the three overall rating measures until we receive that testing data, the reliability data.

The rest of the criteria will focus on the three measures based on the multi-items, so we'll continue on to validity now.

CO-CHAIR PARTRIDGE: Karen, I don't want to slow you down.

DR. PACE: Okay. No, go ahead.

CO-CHAIR PARTRIDGE: But I think, just to be fair to our developers in particular, if anybody has some questions about the validity issues or other issues relating to those three global measures we have deferred and there is information they

might bring back to us during the comment period, I think we should try to raise it now.

For example, are you comfortable with how the global measures are rated, the top box method of scoring and so on? I'm not saying there are issues here necessarily. I just wanted to be sure that if there were, we raise them at this point. Sam.

MEMBER BIERNER: So let me just ask you if you would like to comment on any aspects of that validity testing and any issues that you're going to bring forth, any other information that you were wanting to consider?

DR. PACE: Are you asking about the -- which measures? The three that --

MEMBER BIERNER: The three that were under consideration, because we're not talking about the other three.

DR. PACE: Okay, all right.

MEMBER BIERNER: Or do you want to talk about the three that were not --

CO-CHAIR PARTRIDGE: We're not worrying about the three we've deferred. We're focusing solely on the three that are moving forward.

CO-CHAIR MERLINO: So the comments about validity were for all six. We didn't break them out in the staff comments, did we?

DR. PACE: Right. So and maybe Dawn has something to say about this, but we'll ask the developers to specifically address validity. Because it's hard to discern from this actual validity testing for those three measures. But, Dawn, can you be --

MEMBER DOWDING: That was actually what I was going to ask. I wondered if they could talk us through how they've done the validity testing, because I couldn't actually see it. It was just a repeat of the reliability testing, which I think is what's come up in the notes that we got.

So, yeah. I understand it's pilot

data and it's a pilot instrument, but is there any data on validity?

MS. CRAWLEY: You're correct. We did repeat the information that was in the reliability section in the validity section as well, because it was sort of a combination of the analysis that was done from the pilot data.

We did bring forth today some information that we could share that looked at the composite scores versus the global ratings, to look at the validity of those scores, which were all significant at the 0.001 level. But, again, we did not have it in the packet.

MS. GOLDSTEIN: So we could go through these quarterly. There aren't that many numbers. For nephrologist communications, the first --

DR. PACE: Before you give us the numbers, could you just give us a little more explanation of what numbers you're going to be

giving us? Is it correlation of the --

MS. GOLDSTEIN: It's a correlation between the scales or multi-item measures and the three global ratings.

DR. PACE: Okay, and is this patient level or performance score level or both?

MS. GOLDSTEIN: This is patient level.

DR. PACE: Patient level, okay.

MEMBER BIERNER: We have a white paper board we could actually write --

MS. GOLDSTEIN: Oh, do you want us to do it there?

(Pause.)

MEMBER DOWDING: Sorry. Can we just clarify? We were just discussing. So to meet the criteria for NQF endorsement we have to have validity data at the organization level as well as individual patient level. Is that right?

DR. PACE: Yes, at both levels.

MEMBER DOWDING: At both levels.

DR. PACE: Now, at the organization level it doesn't always have to be empirical. We do still allow for some face validity.

But generally, for the CAHPS measures, they've been kind of doing the same correlation but doing it at the patient level and then also doing it with the computed score level, but I don't know if they've done that in this case.

(Pause.)

DR. PACE: Liz or Barbara, do you want to explain or just --

MS. CRAWLEY: Okay, this was our evidence of construct validity where we looked at the three measures that were made up of multiple items: doctor communication, quality of the center, care and operations and providing information to the patients versus the global ratings of the nephrologist, the staff and the center. And all of these

correlations were significant at p less than 0.001.

DR. PACE: And that's a typical way of looking at validity. So basically the idea is that the scale, you know, it's kind of their hypothesis, that if people are rating doctor communication high, then you would expect the global rating of doctor communication to be in the same direction and, you know, pretty well correlated.

It's not the exact same measure so, you know, you're not going to see, like, 0.9 or 1, but they should be decent correlations and significant.

MEMBER BIERNER: I just wanted to ask, on the third of those, providing information, that's apparently somewhat lower than all the others, what is your interpretation of that and was there anything done as a result of that?

MS. CRAWLEY: We probably have dropped some questions that were on the pilot

testing because of this finding. Originally, the questionnaire was, let me think, 70 questions long.

Now we have it down to 58 because we did drop questions because they did not meet a certain standard that was set. So that might explain why it's so low here. It's still significant, statistically significant.

MEMBER BEVANS: I guess in response to that, my question would be are we being asked to vote on the pilot test version, for which we have evidence, or the sort of revised version of the instrument at this point? Because we don't have those data, right, like, what these validity or reliability statistics look like for the final scale compositions?

MS. CRAWLEY: Not at this time we don't.

MS. GOLDSTEIN: We could.

MS. CRAWLEY: We could, yeah. We could, yeah.

CO-CHAIR MERLINO: You're going to roll out to 5,800 centers? That's the plan?

MS. CRAWLEY: That's the plan.

CO-CHAIR MERLINO: With the pilot from 32 facilities?

MS. CRAWLEY: We have 22 approved vendors who are taking on this task.

MS. GOLDSTEIN: I mean, in addition to the pilot we did, we did do a quality improvement project. So a number of facilities used the survey to make improvements in the operations of the facilities, and it was very successful.

So many facilities have been using, you know, the survey and have been making, you know, quality improvement actions to improve their care.

CO-CHAIR PARTRIDGE: Lisa.

MEMBER MORRISE: I really want to be respectful here and say that I appreciate the effort that has gone into this and feel very strongly about the importance and need of

this level of score.

But I'm feeling also very much like we're not voting with sufficient information based on what you are suggesting is your current effort, and that perhaps what we need to do is go back to the drawing board and look at this again with a more complete set of data, both for us and for those who are not currently in the room.

CO-CHAIR PARTRIDGE: Becky.

MEMBER BRADLEY: There were some questions about the definition of the term "respect" on the survey and how you all defined that. And there was some missing data in some of the tables there related to the respect question. Can you clarify that for us?

MS. CRAWLEY: I think, if I recall correctly, there was one row that had missing information and I think when I looked back at the data that we had on hand, the tables I had on hand, one of the variables got repeated in

the output.

And so we just don't have that at this time. Maybe that's one of the other ones we can ask them to provide for us shortly. I'm trying to remember which one it was. Here it is. The "show respect" under quality, that's the one you're referring to?

DR. PACE: But I think in general and -- Liz and Barbara, you can correct me if I'm wrong -- but in general when that question is asked about are you treated with respect, it's from each individual's perspective of their idea of respect.

So I don't believe we're ever going to see, you know, "and we define respect this way. But correct me if I'm wrong." Is that a fair --

MS. CRAWLEY: Yes.

MS. GOLDSTEIN: Yes, that is correct. I mean, also these correlations when you're looking at them, providing information is not going to be as correlated to the

doctor, the staff, even the center. It's not a direct. It's a little bit different measure.

If they were as high as the quality of the center-doctor communication, you know, I'd be extremely surprised, so.

DR. PACE: Right, and so it's not just about the significance but the size. And what Liz is saying is that from a conceptual standpoint you would not expect that to be as high as the others.

And I think that makes sense and this is at the patient level data but we still, I assume you could do this at the score level as well since you have that data perhaps.

CO-CHAIR PARTRIDGE: Sherri.

MEMBER LOEB: Just a quick, I mean, no, you can't necessarily define respect because it's different for each person. But you have data in other categories for respect, I mean, you have it under nephrologist

communication and caring but you don't have it under the other.

And I would tend to agree with Lisa, that I feel like we're trying to approve something where there's just so many holes that we don't have all the information to really make, I mean, an educated, fair decision.

DR. PACE: Okay, so what we'll do is vote on validity and we're talking about the three measures that you approved under reliability and your options will be high, moderate, low, or if you feel there's insufficient information that's what you would vote here. Does that make sense?

MS. ALLEN: So we'll start the voting on validity. Results for validity, two high, five moderate, three low, seven insufficient.

(Off microphone comments)

DR. PACE: Okay, we're trying to compute here. In order for it to move --

Okay, if the high and moderate is in the 40 to 60 percent range, we consider that the gray zone and we will continue looking at the rest of the criteria.

And then you'll vote on, well, so we'll continue to move through and look at feasibility and usability for the three measures unless anyone wants to make any comment about this vote because it's quite a range from high, meaning there was testing at both levels, to insufficient information, so.

CO-CHAIR PARTRIDGE: I think Liz has dropped off the phone.

MS. DORIAN: Yes, I'm just checking with her. She wasn't there.

DR. PACE: Okay, let's re-vote because Liz has joined us back. So let's go ahead and do re-vote on validity, and I apologize that we didn't get that in, so.

MS. ALLEN: So now we're voting on validity, one high, two moderate, three low, four insufficient. Voting starts now. And

the results are in, two high, six moderate, two low, eight insufficient.

DR. PACE: Okay, still in the gray zone. Still the gray zone so we'll continue through the rest of the criteria for these three measures. Okay, so next is feasibility. And for those of you who reviewed that or any committee members, any comments?

MEMBER BIERNER: I wanted to ask a question. Given, I know this is being done by vendors. Approximately how long does it take to administer? Do you have any feel for that?

MS. CRAWLEY: We have no idea of how long it takes to do it by mail of course. By telephone we are thinking it's an average of about 15 minutes.

MEMBER BIERNER: Fifty?

MS. GOLDSTEIN: Fifteen.

MS. CRAWLEY: Fifteen, 1-5, 15. Well, there are some skip questions. If you, you know, answer some yes or no, then you might skip a block of questions.

(Off microphone comments)

MS. CRAWLEY: We have shortened the telephone script for the intro for the national implementation. I would have to go back and look at what the telephone script looked like.

It was much longer because for the national implementation we, you know, we kept it with few modifications. But, you know, we have shortened it for the national implementation. So, yes, there is an intro in the telephone script.

And in the mail version you get a pre-notification letter and then you get a mail cover letter letting you know the background and why we encourage you to participate.

MEMBER THOMAS: In terms of the feasibility and usability, it strikes me that this is going to be implemented nationwide, you know, this year, next year. It strikes me that it would be really high, that feasibility

would be.

FEMALE PARTICIPANT: The question is moot.

MEMBER THOMAS: Right. I mean, it's going to be in every dialysis clinic across the country soon, right?

DR. PACE: Right, but that doesn't -- Yes, I mean, the question is how feasible you think it is or what the burden's going to be to both patients and facilities, but you're right the --

MEMBER THOMAS: They'll then be tying it to payment and eventually, it's just reporting first and then payment eventually. So not participating is going to be a financial penalty so I would suggest that this is a pretty high feasibility and usability score.

MEMBER BRADLEY: I guess my question sort of relates to that in that earlier in the information it said there were no licensing fees or anything.

But there will be vendor fees associated with this that will be somewhat of an added burden to hospitals that are not currently collecting this data. Is that correct?

MS. CRAWLEY: Each dialysis facility will have to hire a third-party vendor to participate so, yes, there will be those fees that they will have to carry.

CO-CHAIR MERLINO: Any other questions or comments? We're voting on feasibility.

MS. ALLEN: We're voting on feasibility, one high, two moderate, three low, four insufficient. Voting starts now. The results are in, seven high, ten moderate, one low, zero insufficient.

DR. PACE: Okay, let's move on to usability and use.

CO-CHAIR MERLINO: Comments? Questions?

DR. PACE: Everybody knows it's

planned for use, so.

CO-CHAIR MERLINO: Anything? Should we just move to a vote? Peter, do you have a question? No. All right, we're ready to vote. You want to move on, don't you? All right, usability and use, we're voting.

MS. ALLEN: All right, beginning the vote on usability and use, one high, two moderate, three low, and four insufficient information. Starting now. All the votes are in, nine high, five moderate, two low, and two insufficient information.

DR. PACE: Okay, let's move on to the overall suitability for endorsement. And we're talking about the three performance measures based on the multi-item scales in the in-center dialysis survey. Okay.

MS. ALLEN: We're voting on the overall suitability for endorsement for Measure 0258 and we're only voting on those three measures identified, one yes, two no. Voting starts now. All votes are in. Ten

yes, eight no.

DR. PACE: Okay, it remains in the gray zone and probably they'll have some additional information for us on the performance measure level so I think we'll be able to get that cleared up and we will work with them to get the information during the comment period, so thank you.

MS. DORIAN: Thank you. So we'll move on to our next measure which is 0517, the CAHPS Home Health Care Survey. And, Liz, I believe you're staying up here and I'm not sure who else.

MS. GOLDSTEIN: I'm staying here and Lori Teichman is on the phone also for this one.

MS. DORIAN: Hi, Lori. Are you there?

MS. TEICHMAN: Hi.

MS. DORIAN: Hi.

MS. TEICHMAN: Thank you.

MS. GOLDSTEIN: Okay, so I'm going

to give you some background on the survey. This is our Home Health Care CAHPS Survey. This was initially NQF endorsed in March 2009. This survey is required for our Medicare-certified home health agencies and they've been collecting it for a couple years.

For this survey, again, we have very detailed protocols. So this is our manual, pretty thick.

There are approximately 9,000 home health agencies so the results when we present reliability and other information, it's about 1.4 million surveys.

I just want to make one point about the home health industry. There is lots of flux in this industry with agencies coming and going.

So, for example, for looking at improvements over time, we've seen a little movement, not great movement yet, but we think part of this is because we have a number of agencies that close each year and a number of

agencies that come in.

Out of any of our surveys that we've nationally implemented, I think this is the one where we see the most changes in terms of the providers included in the survey.

Just to make a couple more points. I think this survey is unique in that 61 percent of patients responding to the survey are 75 years or older and almost 30 percent are over 85 years old.

So it's definitely an older population than some of our other surveys. It, again, is a very vulnerable population also with, you know, home health workers coming into people's homes.

For this survey, we do have translations in Spanish, Russian, Chinese, both simplified and traditional, and Vietnamese.

Just a couple more points to emphasize is that there are five measures, so there are three measures that are multi-item

and then two global ratings.

And I guess the other point to emphasize, and maybe it's when we get to that section, we have done some analysis of disparities and so I could maybe, I'll hold that information until we get to that section because we do see some disparities in care in terms of this survey. So I think I'll stop there.

I guess another point, and I'm not sure if some of our attachments weren't attached or not because from some of the questions we got from NQF in the Excel file it seemed like all the information that was provided wasn't being seen so I don't know if there was an issue there.

So I'll try, if there are gaps or something, I'll try to fill that in but we may need to have a discussion afterwards and make sure, did you get all the various attachments because some of the stuff we were surprised, that it's right there so, yes, we're not sure.

DR. PACE: Okay, well, and definitely point out something that you think is there and --

MS. GOLDSTEIN: And may not be there or something.

DR. PACE: Well, and it may be. We may have missed it.

MS. GOLDSTEIN: Okay, okay.

CO-CHAIR MERLINO: Any comments? Questions? Evidence. Who reviewed it? Anybody want to kick it off from the phone?

MEMBER BIERNER: I mean, I think there's fairly strong evidence of the need for it. It's very widespread, you know, form of health care. It's nationwide.

MEMBER MORRISE: We utilized home health care in home every night for over ten years and that was pediatric so it would be excluded from the survey.

And I would suggest that eventually pediatric should be included because amongst my group, and I help

facilitate a family group of over 1,000 families in my state, many of whom utilize home health care on a regular basis, they scream all the time about their agencies and whether they are doing a good job or not.

But we have also had a recent encounter as an adult and as I was reading through this I have to say that one of the drums that I beat regularly is that medical reconciliation is more than going over a list of meds.

And I was so excited to see the questions around your prescription medications going in more depth because they did not with us.

CO-CHAIR MERLINO: Liz, any comment on including peds at some point?

MS. GOLDSTEIN: We get that question a lot for a lot of our surveys. Most of our surveys focus, that at least my group works on, on the Medicare population so there aren't too many, you know, pediatric patients

that are Medicare eligible.

That said, I think for some of our surveys we are starting to have some discussions about that.

CO-CHAIR MERLINO: Esther. Dawn, sorry.

MEMBER DOWDING: It's okay. Yes, I just wanted to say coming from a home health care agency I think that this is actually a measure that's used and it's a measure that informs quality improvement programs. So I would strongly support the use of it in terms of the evidence base for it.

CO-CHAIR MERLINO: Chris.

FEMALE PARTICIPANT: And you can probably speak directly to the actions that --

MEMBER DOWDING: Oh, yes. Oh, yes. We take the HCAHPS Home Care Patient Satisfaction Survey very seriously and it is informing quality improvement programs in our agency because of the measures that are coming out. So I would say there's very strong

evidence for it.

CO-CHAIR MERLINO: Chris, and then we'll go to --

MEMBER STILLE: Yes, I just wanted to echo Lisa's concern. You know, for a lot of the measures there's a reasonable rationale for excluding children from this but I would think that the perceptions of parents are pretty much the same as the perceptions of any other caregivers or of other patients, so I really have trouble with the exclusion of children in this.

CO-CHAIR MERLINO: Sherri, then Lisa.

MEMBER LOEB: Taking off my NQF hat and my nurse's hat, from a spouse hat the home care that we received and the medication reconciliation and review of the whole person was by far so exceeded any care we got from any other caregiver that it's truly is a phenomenal thing and I can't speak highly enough and I know all home health agencies

don't benefit that but the home health nurse that we had, he truly made a difference in his care so kudos.

CO-CHAIR MERLINO: Lisa.

MEMBER MORRISE: Yes, the nurses can make and other staff that come from home health can make a huge difference in keeping a patient out of the hospital, which was certainly our case and why we were a part of an approved Medicaid waiver.

And, in fact, most of the children, I don't know if this is true in Colorado, Chris, but most of the children who are able to access home care indeed are Medicaid patients so I think it would be something that CMS would be very interested in, to look at the pediatric population.

And many of those individuals who are receiving service are receiving service from agencies that don't have a pediatric-specific arm but are just home health agencies with a major adult provider cohort as well.

CO-CHAIR MERLINO: Chris, any --

MEMBER STILLE: Right. No, your speculation is right. You know, most kids who would get home health care are receiving SSI and Medicaid and so this is a tremendously important thing to government payers.

CO-CHAIR MERLINO: Anybody want to speak negatively? Why don't we vote on this and then we can move to the next one.

DR. PACE: So we're talking about evidence, that there's rationale that supports that there are health care interventions that can influence the home health experience being measured. Okay.

MS. ALLEN: We're voting on evidence, one yes, two no. Voting starts now. All votes are in. Results 18 yes, zero no.

CO-CHAIR MERLINO: Okay, we'll move on to performance gap.

DR. PACE: Okay, Liz, I think you had some information here. I think there was mention that it was in the measure testing

attachment and I didn't see it there but --

MS. GOLDSTEIN: So I think in terms of the performance gap --

CO-CHAIR MERLINO: Mic.

MS. GOLDSTEIN: Okay, in terms of the performance gap, we have in the form information about what scores, you know, look like today.

But to add to that, we have looked at disparities in terms of these measures. And what we're seeing in particular for the care of patients and the communication between providers and patients, we are seeing within agency effect.

So minorities for these two measures are reporting lower scores than whites. You know, it's not by a huge amount but there are, you know, differences and these are, we're seeing it really within agencies. When we go across agencies versus within, that's where we're seeing it.

So there is clearly, I mean,

general scores are fairly high for all of our different measures but there is variation among the different agencies around the country.

CO-CHAIR MERLINO: Comments? Lisa.

MEMBER MORRISE: So my question to that specifically would be is there a correlation in, so do you look at socioeconomic, although I probably should have drilled down into this information, but is it a socioeconomic situation or a ethnic disparity situation?

I say that because the parents most likely to complain about being treated poorly are those who have the lower education levels and the lower economic level.

MS. GOLDSTEIN: Right. When we're looking at race, ethnicity, we're controlling for the education. And education is our, you know, measure that we have of socioeconomic status. So even controlling for, you know,

education you're seeing some of these differences.

DR. PACE: And just to clarify, Liz, I think the details about ethnicity and race were put into the data dictionary, lots of tables. And so we were just kind of looking for some summary, but that part was in the data dictionary so that was there.

MEMBER THOMAS: But a fair amount wasn't there, right? I mean, I'm looking at some of the staff notes that note that there's no information on the performance provided, no information on the agency performance provided, disparities we just mentioned, no summary of the performance based on performance measures and recommended that this might be considered insufficient. So is some of that covered in the documents that you've said you've submitted that --

MS. GOLDSTEIN: Yes, so the scores are included in the documents we submitted.

DR. PACE: So can you tell us

where?

MS. GOLDSTEIN: So it's in --

(Off microphone comments)

MS. GOLDSTEIN: Yes, measured testing form.

DR. PACE: So we'll pull that up.

MS. GOLDSTEIN: Is the measure patient mix adjusted? Yes, it is. So for --

(Off microphone comments)

DR. PACE: So are you opening the measure testing form so we can see that? So go to 2A. Find 2A and then click on the measure testing. There we go.

Okay, so we have the demographic characteristics in a table, patient mix adjustment factors, factor loading, item total correlations and missing data but --

MS. GOLDSTEIN: There should be a Table 8.

DR. PACE: Okay, Table 8. Thank you. Keep going. There we go. All right. So, Liz, you want to just kind of walk us

through this? Thanks.

MS. GOLDSTEIN: So this is, gives the average median, mode, all the information, standard deviation. These are, you know, agencies that have more than ten completes on the survey. So it's about 8,000 agencies for our five measures.

DR. PACE: And just to explain, the interquartile range is between the 25th percentile of scores and of agency scores in the 75th percentile.

So that means, for example, on care of patients there's a six-point difference between agency scores at the 25th versus the 75th percentile.

Okay. So that's the agency performance on these measures as specified. Okay, so it shows some variability. And any more questions about that or, Lee.

CO-CHAIR PARTRIDGE: Can we go back a couple of minutes in which we were talking about case mix or patient mix

adjustment?

In these numbers that we're looking at, Liz, have you made some adjustments based on the data that you had in the preceding page, that is education or --

MS. GOLDSTEIN: Yes, so for the CAHPS surveys, we always do patient mix. There's been analyses for years and years that certain patient characteristics influence how patients respond to patient experience of survey.

So any data that CMS uses, we do adjust for patient mix. Otherwise, you know, we clearly wouldn't be doing apples to apples comparison. So, again, you know, that is part of our protocols so --

CO-CHAIR PARTRIDGE: But suppose I wanted to know the extent to which this number has been adjusted based on whether or not the population responding is over 75 or completed less than a high school education or so on.

How would I know what the

unadjusted number would have been and do -- I'm just trying to get, I know this matters in some other CAHPS surveys and I'm not sure it matters quite as much here but I think we're going to end up spending some time on it probably later today or tomorrow.

To the purchaser or did the consumer or did the provider, they don't know how much that number has been adjusted by those demographics, right?

MS. GOLDSTEIN: So the providers do know because we always give the models back to them and some of our providers -- I do a lot of stuff on health plans. Those providers know those adjustments by, you know, the exact amount.

CO-CHAIR PARTRIDGE: Yes, I know.

MS. GOLDSTEIN: They have them memorized. So providers do know. We make that information available to them.

The consumer doesn't know that, you know, they don't get when they're

reviewing the website into that detail that the data is adjusted.

But it does ensure, for example, an agency that may be serving, you know, a very different population than your average population, they could, if we don't adjust, they could be hurt by that if they do have a different, you know, group that gives lower responses.

DR. PACE: And let me just clarify that, you know, the NQF criteria do not, I mean, we are endorsing the performance score and, you know, if adjustment is necessary and justified in the analysis it should be done.

So we don't require additional reporting so it is something that, it's not evident from an adjusted score. You're right.

MS. TEICHMAN: I'm sorry to interrupt. This is Lori and I just wanted to let you know that we, you know, we post the data quarterly on Medicare.gov on Home Health Compare and every quarter the patient mix is

recalculated and we post all of the updates to the patient mix adjustments.

And what we do on Home Health Compare is that there's a link so if somebody does want to see what the adjustment scores are, there's a link right to the homehealthcahps.org website and there's a whole table that shows the adjusted factors for the data that was just put up and also, like, a two-page explanation of what it all means.

CO-CHAIR PARTRIDGE: But if I'm trying to figure out whether this home health agency is likely to provide very good care to my grandmother who has only an eighth-grade education, what I'm struggling with is I don't know to what extent these numbers are masking-- information would be useful to me.

DR. PACE: Right, right. So let me just say in general I think, as Liz was saying, the purpose of these are comparative performance assessment.

So the question is, you know, if you're going to do a list and you want to ask the question how would these agencies compare if they were taking care of the same mix of patients, that's the purpose for adjustment.

And so that's all we're asking for at this point, but it's the central question to adjusting for sociodemographic status factors that --

CO-CHAIR PARTRIDGE: And as Karen knows, those of us on CSAC have spent four months on that issue so it's very relevant in our minds.

DR. PACE: Okay, so this is the distribution of the performance score as it's computed with adjustment, and are people ready to vote on performance gap?

MS. ALLEN: So we're starting the vote on performance gap, one high, two moderate, three low, four insufficient. Starting the votes now. All votes are in, 12 high, six moderate, zero low, zero

insufficient.

DR. PACE: Okay. Let's move on to high priority and any discussion here? Everybody ready to vote? Let's move on. High priority.

MS. ALLEN: We're starting the vote on high priority, one high, two moderate, three low, four insufficient. Starting the votes now. All votes are in, 18 high, zero moderate, zero low, zero insufficient.

DR. PACE: Okay, so let's move on to reliability and, again, this would include the specifications including the case mix adjustment and exclusions, et cetera, as well as the reliability testing which we expect at both the patient level instrument and scales as well as the computed performance measures.

And remember, we're looking at five measures so if there are any questions about individual measures to note that and I'll stop there and see what the reviewers --

CO-CHAIR MERLINO: Anybody?

Peter.

MEMBER BIERNER: Oh, I thought it was strong.

CO-CHAIR MERLINO: Peter.

MEMBER THOMAS: So I just want to confirm that 1.5 million patients were surveyed in all states? Is that correct?

MS. GOLDSTEIN: Correct.

MS. TEICHMAN: Yes.

MEMBER THOMAS: That's 10,000 home health agencies annually will --

MS. GOLDSTEIN: This includes about 8,000 home health agencies, the analysis.

MEMBER THOMAS: Any reason why, that you can tell, anything relevant about why 63 percent of the survey takers were women and 37 male? Is that just the demographics?

MS. GOLDSTEIN: I think it's the demographics of who's using home health.

MEMBER THOMAS: Yes. So the algorithm rated this as high? You could rate

it as high, right, 0.7 percent for all performance --

DR. PACE: Eligible. If you had concerns, you could rate it moderate but it was testing at both levels and the data were good. We could bring this up in the measure testing form. I think they have a couple -- It's a little lost in here where we have the --

MEMBER THOMAS: But for all performance measures it was --

DR. PACE: Oh, right. I'm sorry.

MEMBER THOMAS: -- over 0.7?

DR. PACE: There it is.

MEMBER THOMAS: 0.7, right, Cronbach's alpha?

DR. PACE: Right. The Cronbach's alphas are for the three multi-item scales. We don't have the patient level data for the global items but they do have the performance measure level information for all five which you'll see there, .85, .80, .84, .79 and .77.

So those were all above .70. Okay. Any other comments? Okay, ready to vote, reliability for all five.

MS. ALLEN: We're voting on reliability, one high, two moderate, three low, four insufficient. Starting the votes now. All votes are in. The results show 14 high, four moderate, zero low, zero insufficient.

DR. PACE: Validity, so any thoughts about validity testing? Peter.

MEMBER THOMAS: Please, go ahead.

MEMBER LEVINE: I raised this in the work group call. I'm just a little concerned about the standard for inclusion, that you only had to have one home care visit in two months.

It seemed very modest on which to base a judgment as a patient on the home care agency compared to, say, we didn't talk about this but the dialysis center which said you had to have three months of practically daily

experience. So I just wondered how that got decided.

MS. GOLDSTEIN: So, it's --

MS. TEICHMAN: Hi, this is -- Oh, I'm sorry.

MS. GOLDSTEIN: Oh, that's okay, Lori. I was just going to say it's two skilled visits over that two-month period, so one during the sample month and at least one, you know, during the sample month or the prior month.

There are a lot of home health care patients, we'd be excluding a sizable number, that receive maybe an IV or something only once a month.

So we would be -- And a lot of these patients receive home health care for months and months and months or years maybe. So we would be excluding, I can't say the exact percentage but a significant number of patients from our evaluations.

So that's why when the survey was

developed we did not want to, you know, exclude this group. They were having enough experience with the agency in particular since a lot of these patients are receiving it for extended periods of time.

MEMBER LEVINE: I just think for all of those submissions, I think it would really be helpful if that kind of question got anticipated and explained so that we would know why certain choices were made because the questions will always come up.

MEMBER THOMAS: I can't --

FEMALE PARTICIPANT: Oh, I think the phone cut out. No.

MEMBER THOMAS: Some of the notes in the staff notes and then I corroborated it with my own review and I tried to figure out what I thought independently of it and it just strikes me that there are some real questions about whether there is sufficient validity testing here.

So I'm just trying to get a better

sense for, for instance, some of the notes that I was provided with by staff suggested that some of the things that were being put forward as valid to assess validity actually weren't really on target. So can someone explain kind of that perspective because I'm not hearing that come out in our discussion?

DR. PACE: So, Liz, what was presented was item to total correlations and IRT parameters for the home health care CAHPS and generally, a lot of times, the item to total correlations are used as part of the reliability versus validity.

We did find the validity testing in Table, for the measure score, I believe in Table 9, the correlations there. But maybe you want to speak to the item to total correlation.

MS. GOLDSTEIN: I guess it would be helpful to know what other types of things you'd be looking under validity for. I mean, we have data so we can clearly run anything.

DR. PACE: Right. Well, I think what you were presenting before, you were looking at the correlation of the composite to the global rating.

MS. GOLDSTEIN: Which we did.

DR. PACE: This is the performance measure level, right? These are the performance measure -- Yes. So this would be what we are looking for for the performance measure. Exactly.

So the question is were you presenting the, let's see, you had a table of factor loadings and then you had a Table 4 of item total correlations. What were you presenting --

MS. GOLDSTEIN: I mean, I think all of it was presented to look at different aspects of validity.

So, I mean, the item total correlations, we're making sure that the individual items were correlated, you know, within their composite and not to items in

other composite.

So I think we tried to look at it different ways. I think on the form it's not completely clear. It wasn't at least clear to us, you know, exactly what you're looking for for validity.

DR. PACE: Right. I think for some of the CAHPS though measures you've been submitting the, or maybe other CAHPS, but what you showed for the previous one, where you were looking at more the computed composite score correlated to a global rating --

MS. GOLDSTEIN: Yes, well, that's there too. It's here also.

DR. PACE: So, okay.

MS. GOLDSTEIN: That was Table 9.

DR. PACE: Well, that was at the performance score level, right?

MS. GOLDSTEIN: Yes.

DR. PACE: Okay, so in some of the submissions that's been done at both levels, right?

MS. GOLDSTEIN: Yes.

DR. PACE: Okay. So if you could just, you know, I guess tell the committee how you looked at validity of the patient level scales, then they can go ahead and vote. Dawn, Len, any comments about --

MEMBER DOWDING: No, I mean, I didn't have any concerns about this measure but that's probably because I'm very familiar with it, so.

MEMBER BIERNER: Well, I mean, in the Table 8 they give some agency-level descriptive statistics that show certainly a range of responses, some of them very low if you look over the minimal and max portions of the table. So I don't have concerns with it in terms of validity.

DR. PACE: Right, and that's really about the performance spread but that's good. So I'll just ask if you want to make any comment about the -- Okay.

So basically if you look at the

item, the total correlations in Table 4, this is about now individual questions, how it relates to the composite.

So it's a little bit different than what we were looking at before and with some other measures where they actually took the composite and correlated it with some other measure and looked at how they thought that performed.

So this is kind of a deeper level that's often used in constructing the scales in terms of looking for consistency of the item to the scale. So we can --

CO-CHAIR PARTRIDGE: The question before us is those of --

MS. TEICHMAN: This is Lori. I guess I'm at a disadvantage because I'm not in the room and I apologize.

I'll be willing to do whatever it takes so I guess specifically I'd like to know what statistic should be run on what for it to get the answers about the patient-level

validity? What would you like to see?

DR. PACE: So we just wanted you to explain how what you provided is a, how you see that as validity of the scale and we'll have the committee vote on it. And if there's anything else that's needed, we can work with you offline on that but --

MS. TEICHMAN: Okay, fine. That sounds great.

MEMBER LEVINE: Okay. One more. This is an aside but it's getting a little late and I'm getting a little tired.

In reading the questions and I really appreciate when I could just actually see the questions as opposed to we cover five areas and I don't know what they're saying. You know, I was surprised by what's not there.

In my fairly extensive experience with families and patients, mostly Medicare but some Medicaid, almost all elderly, there's nothing in here about how long am I going to get this service because so much of Medicare

is very time limited.

Yes, there are ways of getting an injection, you know, once every week or month or something but most people don't get that, and that's the question they want to know and very frequently the agency doesn't tell them that till the last day and we've documented that.

And the other thing is am I going to have to pay for any of this? Nothing about money. You know, all this equipment, is that going to come? What do I have to pay for? That's what people want to know about and that's what, you know, seems to me is missing from this survey.

MEMBER DOWDING: There is one question on here about level of service, Question 2. "When you first started getting home health care from this agency, did someone from the agency tell you what care and services you would get?"

So we expect them to go through

how many visits and that's actually what it's expected to be. Oh, and actually that's what this question picks up, is that it doesn't happen. Yes.

DR. PACE: Microphone, please. Would you use your microphone please, Carol?

MEMBER LEVINE: It doesn't include duration. As an ordinary person hearing that question, I don't care what, you know, you may intend it to say that but it doesn't say it.

MEMBER DOWDING: Sometimes we don't know.

DR. PACE: Well, I think Carol has a good point around duration but I think that question does address the gist of what we're getting at.

And in terms of payment, I don't think any of the HCAHPS get to how are you going to pay for this? You know, medical finance is probably one of the very top concerns that patients have but I haven't seen it in the other CAHPS measures either.

So certainly a legitimate concern and maybe something we can look at addressing down the road.

MS. GOLDSTEIN: I know patients are very concerned about what they have to pay, but when you go and talk to patients for all of these surveys it's not something when you ask them, you know, how do you define quality of a home health agency or an ICH facility or a hospital, it's not something that comes to the top of their mind.

So I don't think money actually, in all the focus groups I've seen over probably the last ten years in developing these surveys, it's not usually something that they say, you know, defining, you know, quality in a facility.

And we're really trying to get at information, you know, from the patient perspective. Is this agency providing, you know, good care? Putting aside, clearly there are money issues across, you know, all

agencies.

MEMBER MORRISE: So I think that's maybe something that needs to be addressed specifically because it's definitely one of the top concerns that's being mentioned by patient advocates.

DR. PACE: And I'll just mention that NQF has had some work on affordability and have had a group of, that has a lot of consumers and patient representatives talking about measures related to affordability.

And I think, you know, the question is, you know, I think to actually look at how to develop performance measures in that way versus in a experience with care survey all you could get at is were you informed about different things. And I think they're looking even further than that, are there specific measures about affordability.

MEMBER PARISI: Just want to add for those that are not very familiar with the home care operations, to give you a comfort

level, a lot of these things that you're raising are addressed throughout the regulatory requirements.

And it's probably one of the most heavily regulated industries, is the home health environment, so as it relates to payment and educating the patients and the consumers and the family members about payment sources and required documents that are given, it's very prescribed.

So the CAHPS survey isn't designed to look at that. It's more about the perceptions of care so some of those questions as you're evaluating the responses you may drill down to see what information may or may not have been lacking but it should be addressed in the home health environment.

MEMBER DOWDING: Yes, can I just echo that? We actually ask our patients if the nurses explained to them how the care will be paid for so it's part of our own internal patient satisfaction survey. So most home

health care agencies collect that information. It's just not reported in the HCAHPS.

CO-CHAIR MERLINO: Any other questions, comments? So let's vote.

MS. ALLEN: So we're voting on validity, one high, two moderate, three low, four insufficient. Starting the vote now. All votes are in. Results show 11 high, six moderate, one low, four insufficient, zero, sorry, zero insufficient.

DR. PACE: All right, let's move on. Any comments?

CO-CHAIR MERLINO: Let's vote.

MS. ALLEN: We're voting on feasibility, one high, two moderate, three low, four insufficient. Voting starts now. All votes are in. Results show 17 high, one moderate, zero low, zero insufficient.

DR. PACE: Okay, usability and use. So these are currently publicly reported, right, Liz?

MALE PARTICIPANT: Yes.

DR. PACE: Yes. Right.

MS. GOLDSTEIN: They are publicly reported.

DR. PACE: Okay.

CO-CHAIR MERLINO: Any questions, comments? Let's vote.

MS. ALLEN: We're starting the vote on usability and use, one high, two moderate, three low, four insufficient information. Starting the vote now. All votes are in. Results show 17 high, one moderate, zero low, zero insufficient information.

DR. PACE: Okay, overall suitability for endorsement.

MS. ALLEN: Now we're voting on overall suitability for endorsement for Measure 0517 CAHPS Home Health Care Survey, one yes, two no. Voting starts now. Please vote again. We're still missing two votes. Please vote again.

MALE PARTICIPANT: Everybody check

their green lights. There you go.

MS. ALLEN: Thank you. All votes are in, 18 yes, zero no.

CO-CHAIR MERLINO: So we're going to take a 15-minute break. We're a little ahead of schedule actually so that's good.

DR. PACE: Right, so let's --

CO-CHAIR PARTRIDGE: Oh, I'm sorry. And I think in a way we've saved the best for last because one of the` measures we're going to be discussing involves a new measure, a new-ish measure, CTM measure and how it's going to mesh with HCAHPS and I think this is actually going to be kind of a fun discussion. And Eric Coleman will be calling in from Colorado, yes.

DR. PACE: So let's reconvene at 3:20. Thank you.

(Whereupon, the above-entitled matter went off the record at 3:02 p.m. and resumed at 3:20 p.m.)

MS. DORIAN: Okay, everyone. Just

wanted to call your attention to the slide. If you were interested in dinner, we do still have a block of reservations. And I have the name and the location of the restaurant on this slide, so if you wanted to make a note of that.

A few people stopped over to let me know, so you still have time. You can send me an email or just stop by.

(Off microphone comments)

MS. DORIAN: Mio, yes. It's under my name. That's a good point. Or you could say NQF, National Quality Forum. We're there often.

So shall we move on to the next measure which is 0166: HCAHPS. And we have -- oh, Bill, you're here, great. You want to give an introduction to the measure?

MR. LEHRMAN: Thank you. Hello, I'm William Lehrman. I'm the government task leader for the HCAHPS Survey at CMS. And I'd like to give a few brief remarks about the

HCAHPS Survey. It's the first national standardized, publicly reported survey of patient experience of care.

Development of the HCAHPS Survey began in 2002. The NQF endorsement came in 2005. We nationally implemented it in 2006. HCAHPS results or participation become tied to pay-for-reporting for hospitals beginning in 2007.

We began the public reporting of HCAHPS scores in 2008. NQF and re-endorsement occurred in 2010, and then the linkage of HACHPS scores to pay-for-performance or hospital value based purchasing, the CMS program, that began in 2012.

So once again, HCAHPS is a national, standardized, uniform survey of patient experience. It's widely used. We believe it's becoming more and more influential. It's available in English, Spanish, Chinese, Russian, Vietnamese and Portuguese.

For the scores publicly reported just this month, July, which are based on patients' discharge between October of '12 and September of 2013, there were over three million completed surveys and over 4,000 hospitals participated.

We've seen improvement in HCAHPS scores since the scores were publicly reported. We've seen significant improvement across all ten and soon to be 11 measures in the HCAHPS Survey.

And just to mention what those measures are, they are, briefly, communication with nurses, communication with doctors, responsiveness of the hospital staff, pain management, communication about medicines, cleanliness of the hospital environment, quietness of the hospital environment, discharge information, overall rating of the hospital, would you recommend the hospital, and just added last year and will be publicly reported for the first time in December is a

care and transition measure.

And that is a brief overview of the HCAHPS Survey.

CO-CHAIR MERLINO: Any questions on evidence? Very familiar with this one. Comments? Shall we vote and move on? All right, let's vote.

MS. ALLEN: So we're voting on evidence, and one yes, two no. Voting starts now. We're still missing a vote. Okay. All votes are in. Seventeen yeses, zero no.

CO-CHAIR MERLINO: Okay, moving to performance gap. Any comments, questions?

CO-CHAIR PARTRIDGE: I can't remember where these are in the materials submitted to us, but as I have a note, a couple of these are fairly well topped out. We got that table someplace?

DR. PACE: So go to 1(b), 1(b)(2), I think, was some --

(Off microphone comments)

CO-CHAIR PARTRIDGE: Yes.

DR. PACE: So just one comment. It's problematic to enter this kind of data into our submission forms so that's why it's a little hard to follow, but the first number is the mean performance.

So, for example, communication with nurses, 79.12 would be the average performance across hospitals. And then it goes to the 99th percentile, 95th, et cetera, and the very last number, which is on the second line, is 6. That's the interquartile range.

Okay. Just to try to orient people because I know it's kind of hard to follow. So Lee, did you have a specific one that --

CO-CHAIR PARTRIDGE: I'm having trouble reading these numbers at this distance.

Okay, I think what was happening was I was confused by that chart.

DR. PACE: Yes, sorry.

CO-CHAIR PARTRIDGE: I was reading some 100s and getting very concerned. But I'm much less concerned. Thank you, Liz.

MS. GOLDSTEIN: You're welcome.

CO-CHAIR PARTRIDGE: What I'm looking at is some means that in fact show there's plenty of room for improvement.

CO-CHAIR MERLINO: Any other questions about performance gap?

Chris?

MEMBER STILLE: So just in terms of disparities it said that disparities information was calculated but it was in an appendix, and I don't think we have easy access to that. Any sort of rough summary or anything that jumps out?

MR. LEHRMAN: Yes.

MEMBER STILLE: It would just be good to kind of know.

MR. LEHRMAN: We do find racial ethnic disparities on the survey non-Hispanic whites tended to score better than minority

groups. But within the same hospital, minority groups, blacks and Hispanics, tended to score higher than white non-Hispanics.

The issue, really, is minority patients tend to go to poorer performing hospitals on average than white patients. So we still find differences between Asian patients and white patients, Asian patients are more critical about their hospital experience even within the same hospital, but Hispanics and blacks tend to go to hospitals that on average have poor patient experience for all patients.

CO-CHAIR MERLINO: Any other questions? Comments? Okay, should we vote on performance gap?

MS. ALLEN: So we're voting on performance gap. One high, two moderate, three low, four insufficient, starting now. All votes are in. Sixteen high, one moderate, zero low, zero insufficient.

CO-CHAIR MERLINO: High priority,

moving right along. Any comments? Questions? Does anybody think this isn't a high priority? Shall we vote? We might be out of here for lunch.

MS. ALLEN: We're voting on high priority. One high, two moderate, three low, four insufficient. Voting starts now. All votes are in. Eighteen high.

So Liz is on the phone, and at times she's voting and sometimes she's not.

MS. DORIAN: We just missed her vote on the very first one. I didn't get to it in time, but add it in.

MEMBER MORT: I am here. I don't think I've missed any vote. Maybe --

MS. DORIAN: I think just that first one there was a lag and it came in late. But we'll be sure to add it in.

MEMBER MORT: Oh. I was a big fan.

MS. DORIAN: Yes.

CO-CHAIR MERLINO: Okay,

reliability.

DR. PACE: And Liz, feel free to comment, you know, we're interested in your thoughts as well.

MEMBER MORT: Thank you.

CO-CHAIR MERLINO: Moving to reliability. Comments? Questions?

DR. PACE: So under reliability we talk about measure specifications as well as the reliability testing. And --

MS. GOLDSTEIN: I was going to add I put, actually, I probably should move the Cronbach alphas are over there.

MEMBER BRADLEY: I noticed in, I think it was one of the other CAHPS Survey, it mentioned that risk adjustment is not appropriate for these types of measures. It's more case mix. But then you talk about risk adjustment for this CAHPS measure. Is there, kind of what is the philosophy? Do we go with risk adjustment or case mix adjustment or --

MS. GOLDSTEIN: Yes. We always, I

think the form called it risk adjustment, so it may have lapsed into that word. But we always call it case mix or patient mix adjustment particularly in the hospital where it's called patient mix adjustment.

But it's the same for all the CMS surveys. We do adjust for characteristics that influence response tendencies.

MEMBER BRADLEY: So in your research where you show that Asian populations don't tend to score as high, would you adjust for that?

MS. GOLDSTEIN: We don't adjust for race ethnicity. So we adjust, in the hospital world we're adjusting for age, education, overall health status, their service line, whether it's maternity, surgical or medical, and - self-referred health status. Health status.

MEMBER BRADLEY: Okay, thank you.

CO-CHAIR MERLINO: Have you looked at the data yet on depression? Are you

adjusting for that? Do you see any differences by patients who report, self-report depression?

MR. LEHRMAN: We did an item last year to the survey about overall mental or emotional health because of suggestions that perhaps depression or other things weren't being captured by the overall general health. Analysis showed that, which we kind of knew before from our pretesting, but it didn't really add anything to what was explained by overall general health.

So we have looked into that but did not find that patient assessment of their overall mental or emotional health impacted the scores more than could be accounted for by overall general health or any other patient mix adjusters.

CO-CHAIR MERLINO: So didn't change overall self-reported health, right? Health status?

MR. LEHRMAN: It didn't add to the

adjustment.

CO-CHAIR PARTRIDGE: But going back to Becky's question about a case mix adjustment, CMS makes those adjustments. But if you look at the CAHPS' instructions, you'll see that other users can choose to do that or not.

In other words, for your purposes those are a fixed component of these measures. But if Jim decided he wanted to use CAHPS and didn't want to follow the protocol, you wouldn't necessarily have to adjust. If you're reporting it to CMS, you do, and CMS will adjust it. But if you wanted to report it out yourself without the adjustment, you could.

MS. GOLDSTEIN: Yes. So for surveys that if you're using it for quality improvement, so to compare your performance over time in your hospital, say, then it really doesn't need to be adjusted unless your patient mix changes dramatically from time

period to period.

It's when we're using it for public reporting or accountability in comparing hospitals, that's when it needs to be adjusted.

DR. PACE: So let me just also, just to get us back on track. We talk about case mix adjustment under validity, so if we could hold that until we get to validity.

But Bill, if you want to answer that question.

MR. LEHRMAN: Well, I just want to mention that we do report every quarter on our own website, the patient mix adjustment coefficients. So I think you can work backward from those to get the unadjusted scores. We also adjust for survey mode of course, telephone, mail, et cetera, but the hospitals submit the data to us. The data they submit to us is unadjusted, so they have that if that's what they want to work with.

CO-CHAIR MERLINO: Any other --

Carol?

MEMBER LEVINE: Why are patients who go to SNFs excluded? Is it only hospital to home? And are patients under observation status, are they considered to be included?

MS. GOLDSTEIN: So in terms of if they go to a nursing home from the hospital, we actually initially included that group of patients. We have over time excluded that group because response rates were really, really, really low for that group. It was very hard to reach them.

Often the vendors may not have had the right mailing address, the survey never got to them. If they tried to call the nursing home, the nursing home staff would refuse to transfer them to the patient's room. So response rates, vendors were spending, for a small population were spending a very significant time getting nonresponse. So they eventually were excluded from the survey.

MEMBER LEVINE: And a little

disturbing, because so many patients now do go to short term rehab. And that is part of the discharge process is did you know where you were going, you know, and all that. But I understand.

MS. GOLDSTEIN: Observation patients are excluded from the survey.

DR. PACE: So what's up here is the performance measure score reliabilities, and I believe you used the interunit reliability that we mentioned in the notes. And then Liz posted up here, because we didn't have that, the Cronbach's alpha for the patient level scales on the measures that are multi-item.

MEMBER BEVANS: Is the discharge info only two items? Could you explain the --

MS. GOLDSTEIN: That one's just two items, so that's why that score's pretty low.

CO-CHAIR MERLINO: Any other questions?

MEMBER NEUWIRTH: This is just for us. Out of curiosity, is there a SNF type CAHPS that we'll be reviewing?

DR. PACE: There are three nursing home CAHPS. One for short term, one for long term and one for family members. Those surveys and measures are sponsored by AHRQ at this point in time, and those were part of the list of measures that were not submitted for endorsement maintenance and we will have to follow up with the developers in terms of where we stand with those. But there are three.

CO-CHAIR MERLINO: Anybody else?

MEMBER NEUWIRTH: I guess I just wanted to echo what Carol said. I think, you know, more and more with elective surgeries and the elderly population getting more elective surgeries, people are going to continue to, you know, go to these short term facilities for rehab, and we're not going to capture them in this way. So I think that's

a big gap.

MEMBER BRADLEY: Could I just say there's not a CAHPS survey, either, for inpatient rehabilitation hospitals. Are you all working on one of those, for IRFs? A CAHPS survey for IRFs.

MS. GOLDSTEIN: There is discussion about that so I'm not sure, you know, what CMS is going to do. The areas that we're working on right now, besides hospice CAHPS which we just finished, we've been doing some testing of an emergency department survey as well as a surgical outpatient for both hospitals and ambulatory surgery centers. So those are the ones that are closest on the horizon.

CO-CHAIR MERLINO: Anybody else? Peter?

MEMBER THOMAS: As I understand it we will be looking at later in this process, functional measures, and some of those might get to some of the rehab issues, and maybe not

IRF specifically, I don't know. But we'll see what the measures are. But am I right about that?

DR. PACE: Yes. We will in Phase 2 be looking at function measures. It's more, you know, very specific about patient functional status, so it's not experience with care, these type of experience measures from that facility.

CO-CHAIR MERLINO: Any other comments? Questions? Shall we vote?

MS. ALLEN: We're voting on reliability. One high, two moderate, three low, four insufficient. Voting starts now. Please revote. All votes are in. Results, 14 high, 4 moderate, zero low, zero insufficient.

CO-CHAIR MERLINO: Validity. Comments? Any more risk adjustment comments?

MEMBER STILLE: This is relatively small, but I was wondering in your case mix adjustment had you thought at all, maybe this is the next round thing, about looking at

complexity as something you adjust for? There's more sophisticated complexity scores that are going on and people are finding that there's relationships with experience of care.

MS. GOLDSTEIN: There has been some work early on with the pilot looking at different variables and looking at, you know, more condition information and more complex models. It didn't really give us anything more in terms of the adjustment. I think some of that's in the published literature too.

CO-CHAIR MERLINO: Any other comments? Questions?

MEMBER NEUWIRTH: I was confused. The domain area around quiet and physical environment, were those ultimately combined or are they now separated out?

MR. LEHRMAN: For public reporting they're separate. Quietness and cleanliness, we did focus group and other testing, and consumers said that they really thought those were very different aspects of the hospital.

So initially we were thinking about combining them. We decided not to. On the other hand, for the value based purchasing program we do combine them into what we call the hospital environment dimension in VBP.

MEMBER NEUWIRTH: Thank you.

CO-CHAIR MERLINO: There was talk at one point about stratifying by hospital bed size. Are you still thinking about that or has there been any other consideration?

MS. GOLDSTEIN: That isn't, I mean we wouldn't use it in case mix or patient mix, so it's a hospital characteristic. That's, you know, under the control of the hospital. We haven't had more discussions about stratifying on the website. It just makes the displays fairly complex, and it would have to be done for the clinical measures also, although there's been lots of work currently in CMS going on to try to simplify some of this information for consumers.

CO-CHAIR MERLINO: Do you have an

idea of how much traffic comes to the website from consumers? I'm just curious.

MR. LEHRMAN: We're not sure how many consumers go there, but the hit rates are 15,000-20,000 per month, I think, for HCAHPS online. That's our own HCAHPS website.

CO-CHAIR MERLINO: That's Hospital Compare?

MR. LEHRMAN: For Hospital Compare.

MS. GOLDSTEIN: Oh, to get Hospital Compare, you could just put hospitalcompare.gov and --

MR. LEHRMAN: Yes.

CO-CHAIR MERLINO: So it's about 20,000 a month?

MR. LEHRMAN: Yes. Not as many consumers go there as we'd like them to.

CO-CHAIR MERLINO: Any other questions about validity?

DR. PACE: So they submitted some information I think we posted last week in a

supplemental form that relates to the correlations at the patient level and then at the hospital level, so Nadine is pulling those up now.

So this first one was the patient level. So I'll let Liz or Bill just kind of explain what this represents or what we might take from this.

MR. LEHRMAN: Okay, so these are top box correlations at the patient level from patients discharged in the year 2013. The top box is the most positive response in the survey. Typically that's always, nurses always communicated with courtesy and respect, for instance. We find fairly high top box correlations amongst and across the measures.

These are all 11 measures including the new care and transition. I guess we'll be talking more about that later. We're pleased the level of correlations. They're high but not too high.

We think that these different

domains or measures are tapping into different aspects of the patient experience. And a hospital that, the level of correlations is such that we don't believe there's a lot of redundancy or halo effect in the survey.

We also submitted, as you mentioned, a little bit late the hospital level top box correlations which are higher than the patient level ones you see here. And also the average top box inter-item correlations which shows, the bottom table that shows the correlations of the individual items that go into these seven composites.

As you notice, the nurse with nurse is 0.50 indicating that the items within that nurse composite were more highly correlated with each other than they are with, say, the other six. We call them composite measures which we think is also a good thing. Again, these are all significant at about any level you care to measure at.

DR. PACE: Right. So I think that

the two that we're most used to seeing are the ones that we talked about, the patient level and the hospital level correlations of the scores.

And as Bill was saying, the basis of this is that they would expect them to be correlated because they're all tapping in to an overall idea about experience with care. But, you know, obviously they're not redundant of one another.

So then again under validity, if you have concerns about case mix adjustment or exclusions, missing data, that should be in the testing attachment as well. So any questions or comments?

MR. LEHRMAN: I would just like to reiterate as Liz said earlier that we do a lot of training. This is the Quality Assurance Guidelines Version 9.0. This is the ninth revision, or eighth revision of the original.

We take oversight very seriously as well as training the vendors and individual

hospitals' administrative survey, and we found that you almost can't be too specific on how you indicate how things should be done with just about every aspect of the survey to ensure that the results are comparable across hospitals.

There are some 50 approved vendors and over 4,000 hospitals doing the survey. So creation of standardization and oversight and training are very important to the health of the survey and the usefulness of the results.

CO-CHAIR MERLINO: Any other questions or comments? Should we vote on validity?

MS. ALLEN: Voting on validity. One high, two moderate, three low, four insufficient. Voting starts now. All votes are in. Seventeen high, zero moderate, zero low, zero insufficient.

CO-CHAIR MERLINO: Feasibility. Comments? There's a lot of experience with this one. Anybody? Questions? Shall we

vote? Feasibility.

MS. ALLEN: Voting on feasibility, one high, two moderate, three low, four insufficient. Voting starts now. All votes are in. Seventeen high, zero moderate, zero low, zero insufficient.

CO-CHAIR MERLINO: Usability and use. No comments? Questions? Come on.

All right, if there's no comments or questions we'll take a vote.

MS. ALLEN: Voting on usability and use. One high, two moderate, three low, four insufficient information. Voting starts now. All votes are in. Seventeen high, zero moderate, zero low, zero insufficient information.

CO-CHAIR MERLINO: All right, moving on to 0228 3-Item Care Transitions Measure.

(Simultaneous speaking)

CO-CHAIR MERLINO: I almost had us there. All right, any comments overall?

MS. ALLEN: We're voting on overall suitability for endorsement for Measure 0166-HCAHPS. One yes, two no. Voting starts now. All votes are in. Results are 17 yes, zero no.

MR. LEHRMAN: Thank you.

MS. DORIAN: Okay, so now we are up to 0228 which is the CTM-3, 3-Item Care Transition Measure from the University of Colorado. And I believe we have Eric Coleman on the line, correct?

DR. COLEMAN: Hi, yes. Eric Coleman's on the line.

MS. DORIAN: All right, Eric. Welcome.

DR. COLEMAN: Thank you. I wish I could be there in person to meet with you all face-to-face.

MS. DORIAN: Yes. That would be nice.

DR. COLEMAN: Should I just give a quick thumbnail about the Care Transition

Measure 3-Item?

MS. DORIAN: Yes, perfect. Thank you.

DR. COLEMAN: Great. So we have had the privilege of going through the NQF endorsement process twice before. In these cases it was under the hospital care coordination track.

We're very excited to be in the Person-and-Family-Centered Care track because, really, the essence of a care transition measure gets to the extent to which patients feel as though they've been prepared for their self-care. We deliberately spent time in the early and developmental phases to co-design this with consumers to make sure that these items resonated with them.

As you've heard from Bill and Liz, the 3-Item Care Transition Measure has been incorporated into HCAHPS and we really value the spirit with which they have been collaborating with us. But I think I'll just

keep it brief and stop there.

DR. PACE: Eric, do you want to just mention the change that you sent us just so everyone knows where you're at in terms of --

DR. COLEMAN: Sure. Thank you, Karen. So we, as I mentioned, have had good communication back and forth with the HCAHPS folks. And in order to avoid creating confusion in the field we've really tried to align the specifications of the care transition measure with the way it's being used in HCAHPS.

And so the most recent change which came, I think, when Lee Partridge made a very helpful suggestion about a month ago that we should also consider adopting the top box approach to reporting. And we have done that.

We've also reached out to our partners who are using these in a performance measurement, performance reporting context, to

see if they would be willing to go with us in this direction and they're all in agreement.

DR. PACE: Okay, thank you.

And just for those who were on the workgroup call that's important, because some of the testing that was submitted on the performance score level was actually from the HCAHPS testing. And so it needed to match the specifications, and so I think that's in sync now.

CO-CHAIR MERLINO: Anybody from the call want to talk about it?

CO-CHAIR PARTRIDGE: I was on that call. I think there were several of us on the call. This is a measure that the organization that I'm associated with is very, very fond of.

We've done a lot of work around patient safety and readmissions, and we happen to think that the three questions that are asked here are pretty relevant to whether or not a patient reappears on your doorstep, and

also whether or not the care that they need post-hospitalization is provided.

And we did notice that a lot of the work that Eric Coleman and his colleagues have done on the measure and the testing data that we had was different from what would be or might be the results once they were incorporated into CAHPS, the top box score being one of the issues.

Another was, as I remember, Eric, you didn't use some of the case mix adjustments that HCAHPS typically uses in --

DR. COLEMAN: Yes, that's correct.

CO-CHAIR PARTRIDGE: And are you now comfortable with using those case mix adjustments as well in this measure?

DR. COLEMAN: We're certainly very comfortable adopting those approaches. I don't know if Bill and Liz are still on the line, but they did share some of their data with us and so we had an opportunity to see just how much the measures fluctuated under

that kind of adjustment. It's pretty modest. But in the spirit of alignment, we'd be happy to go that route too.

DR. PACE: Right. But for right now what you submitted was just the scoring, not adding the case mix adjustment, correct?

DR. COLEMAN: We have not sent that in yet.

DR. PACE: Right.

DR. COLEMAN: We had understood that it probably would be a source of discussion on this call and we discussed it as a team are very open to that, submitting that as an amendment or how ever you would call it.

CO-CHAIR MERLINO: Anybody else on the call that wants to comment? From an operational standpoint I think this is a very important measure. It's a gap that we need to fill.

DR. PACE: So should we start with evidence then? This is something that hospitals and other providers can influence.

Do people agree, or any comments on that?

CO-CHAIR MERLINO: Chris?

MEMBER STILLE: Just real quickly, from the pediatric perspective it's at least as important if not more important. So I'd encourage a version for child health to be developed as well.

DR. PACE: Eric, are pediatric excluded?

DR. COLEMAN: You know, I'm really glad that question's been brought up. We from the beginning saw this as really expanding across the age spectrum. And again in the interest of being aligned with HCAHPS we did add the 18 and older.

However, in order to try to create balance in the universe we've been working closely with our colleagues in Boston who are trying to create a more pediatric oriented version of this and sharing our materials, literally, with them hoping that they'll one day be before you with a pediatric version.

DR. PACE: And actually, tomorrow we're going to be looking at a child HCAHPS. So I think that's a good question that we can, I didn't think of that before.

CO-CHAIR MERLINO: Carol? Oh, Sherri?

MEMBER LOEB: I agree with Chris. You know, inpatient HCAHPS is really important, but transitions of care, you know, once you're home is what we need to prevent readmissions and keep the patients safe once they're out of the hospital. So it's really vitally important.

CO-CHAIR MERLINO: Any other comments? Can we vote on evidence and then get more granular? All right, let's vote.

MS. ALLEN: We're voting on evidence. One yes, two no. Voting starts now. The results are in. Seventeen yes, zero no.

CO-CHAIR MERLINO: Okay, moving to performance gap. Lee, do you want to kick it

off with a comment?

CO-CHAIR PARTRIDGE: No, I don't have a comment on this one.

CO-CHAIR MERLINO: Anybody? Comments, anybody else on the call? So staff notes? Pretty good?

DR. PACE: Right. So Eric, for the performance score you basically had the information from Maine's utilization. And so I don't know if you want to make any comments about performance gap.

DR. COLEMAN: Sure. I do appreciate the comments that were shared earlier about chances are in our professional lives and in our personal lives we all can point to times when the experience of discharge from the hospital, particularly with the focus on self-care, hasn't gone as well as we would have liked.

Our experience in the state of Maine, and we're also working with CMS under the Community Based Care Transitions Program

that's out of CMMI, you know, we are finding consistently, I guess, the nice way to say it is room for improvement among the earlier HCAHPS testing.

And I don't know, again, if Bill and Liz are still on the line, but among all the subcomponents of HCAHPS the care transition measure scores are the lowest.

And when we look at areas around safety, when we look at areas around quality, and then I think when we really start getting into the nuts and bolts of what the patient experience is like, which is really the thrust of the program that I run in Colorado, I think for the most part that this is an area that really warrants greater attention from the performance standpoint.

CO-CHAIR MERLINO: Any comments or questions? Concerns? I think we can vote.

MS. ALLEN: Voting on performance gap. One high, two moderate, three low, four insufficient. Voting starts now. All votes

are in. Sixteen high, one moderate, zero low, zero insufficient.

CO-CHAIR MERLINO: Okay, moving to high priority, I would just add what Chris says. This isn't just satisfaction, this is safety and quality as well. This is a huge gap coming out of the hospital so it is a high priority. And absolutely, it needs to be in pediatrics. No question.

Any comments? Questions? Shall we vote? Vote on high priority.

MS. ALLEN: Voting on high priority. One high, two moderate, three low, four insufficient. Voting starts now. All votes are in. Seventeen high, zero moderate, zero low, zero insufficient.

CO-CHAIR MERLINO: Okay, reliability. Karen, any comments?

DR. PACE: Eric submitted the patient-level instrument reliability. And I guess, Eric, there were two different results were presented, 0.8 and 0.93. It said that

was for four factors, but it was a little unclear what you were presenting. So maybe we can just go to the testing attachment and you can clarify that for us.

DR. COLEMAN: Yes. Thank you, Karen. And we realize that it's the pros and cons of having been around a little bit that we have in doing different types of testing over time.

As Karen mentioned, our most recent testing looks at Cronbach's alpha 0.80 and a Spearman-Brown reliability of 0.84. That was in collaboration with our HCAHPS colleagues. In some of our other work, we have identified a Cronbach's alpha as high as 0.93.

And one of our, I think just to the question about how the measure performed in more diverse populations, we did the study where we deliberately oversampled people, African Americans, Hispanic American and folks in rural dwelling areas, where we found very

high Cronbach's alpha, in this case 0.94 in African Americans, 0.93 in Hispanic Americans, 0.96 in rural dwelling adults. And so that might explain a little bit of the earlier confusion.

DR. PACE: Okay. Thank you. And can you just clarify what this meant about four factors?

DR. COLEMAN: Yes. There was the three factors each related to the three items and then there was sort of an overarching factor. And I'm sorry, I'm scanning to figure out which page I share that with you. But the reason why there were three items in four areas was that one was sort of a cross-cutting factor, cross-cutting characteristic.

DR. PACE: CTM, what you're presenting for endorsement is three items, correct?

DR. COLEMAN: Three items, yes.

DR. PACE: Okay. All right.

CO-CHAIR MERLINO: Comments?

Questions? Oh, sorry. Dawn?

MEMBER DOWDING: Could I just clarify? And I might have missed this, but am I right in thinking that you've just changed the way in which you score the CTM? So it's gone from an average rating to a top box rating.

So the data that we're looking at, is that being calculated for the top box scoring or is it on the old way of scoring? Because I know there might not be a difference, but it is important for us to understand.

DR. COLEMAN: No, I appreciate it and I'm glad you brought that up because we are a little bit in a transition phase ourselves in terms of the work that we're presenting.

You know, essentially in the earlier days we did present this as a score that was standardized to a 0-100 scale, a linear transformation, if you will. And then

as time has evolved and we've followed where HCAHPS is going, we have tried to align with them using the top box approach. The material that go into both calculations are the same. It's really sort of how the information is being presented.

But the earlier work we did was based on a 0 to 100 linear transformation score, and the more recent work is using the top box approach. And we're very comfortable using that top box approach from here forward to be, as Lee suggested, to avoid creating confusion in the field.

DR. PACE: So let me just clarify. The Cronbach's alpha there is at the patient-level instrument, so that isn't affected by the hospital scoring. And the hospital-level reliability that they presented is based on the top box scoring because it's data that was shared from the HCAHPS data that they presented. Is that correct, Eric?

DR. COLEMAN: Yes. Thank you,

Karen.

DR. PACE: Okay.

CO-CHAIR MERLINO: Any other comments? Dawn, you're good? All right, should we vote on reliability?

MS. ALLEN: Voting on reliability. One high, two moderate, three low, four insufficient. Voting starts now. All votes are in. Results, 13 high, 3 moderate, 1 low, zero insufficient.

CO-CHAIR MERLINO: Okay, validity testing.

DR. PACE: So let's see. Eric, maybe you could tell us how you did validity testing of the patient-level instrument. It's like you -- oh, I see what you did. You compared patients who had ED or rehospitalization to those who hadn't, to look at this is what we call known groups' validity testing.

So you have a group that had ED and rehospitalization visits and you expect

that maybe they had a different care transition experience than those who were not readmitted and what the difference was. So Eric, do you want to say anything more about that?

DR. COLEMAN: Yes. No, that's absolutely right. We thought since again kind of back to first principles, if really the three items are telling us the extent to which hospitals help prepare these folks for what they were going to experience after they left the building, hospitals that did a better job should expect fewer people to have to come back again. And so Karen's right on the dot.

We also did another level of validity where we knew something about the hospitals that we were comparing. So we chose hospitals that were part of a larger umbrella organization for which we already had a fair amount of insight into who does a really good job in this area and who didn't.

And so we were able to then again,

you know, with the added insight of what went on kind of underneath the hood we were also able to do another form of validation and write that up for publication that CTM was able to discriminate among the hospitals in the way that you would have guessed based on knowing which hospitals do a better job in this area and which didn't.

So we had a fair level of insider's information, if you will, about how these hospitals performed then allowed us to step back, test the measure as though we were sort of objective and then look at the results.

DR. PACE: And then for the computed performance score using this top box method, you submitted again the analysis that was from HCAHPS where you looked at the correlation of the care transition measure with the other measures within the HCAHPS set.

DR. COLEMAN: That's correct.

DR. PACE: Okay.

CO-CHAIR MERLINO: Comments or concerns? Estee?

MEMBER NEUWIRTH: This is Estee Neuwirth and I'm a huge fan of transition support. I guess I'm wondering to what extent these measures have been harmonized with the HCAHPS particularly with the discharge section of HCAPHS, which seems so comparable.

So I'm wondering, you know, how distinct are these and what kind of testing has been done to distinguish the power of each of these different elements of the HCAHPS measure and of the CTM measure, and in particular just because I'm wondering about sort of duplicative issues.

DR. COLEMAN: Yes, and thank you. And just to sort of time travel together back a number of years, when HCAHPS first came out we had some concerns about the discharge items within HCAHPS. And because those folks had been so collaborative we also were able to see how those discharge items were performing.

And we had a hunch that our CTM-3 items and the HCAHPS discharge items were measuring different constructs. We did do some head-to-head comparison. I'm happy to share that document with the group.

We didn't go forward to publish it because we, again this was done, really, I think, for us to then have another conversation with the HCAHPS developers about whether CTM might complement the existing structure.

But I will just share the Cliff's notes version is that we were able to demonstrate that the two former discharge planning items are different than the three CTM items when we looked at different forms of validation. And if that's a document that the committee would like to look at I'm happy to share that. I think we submitted this with our earlier endorsement attempt.

DR. PACE: I need to try to pull it up. Do you have offhand the correlation

between the, in HCAHPS, the CTM or the care transition and the discharge performance measures?

DR. COLEMAN: I'm trying to pull them up, and by the way the hold music is not coming from us. Again our goal was not to poke holes into HCAHPS but it was to show that there probably was a different --

MEMBER NEUWIRTH: I'd be interested in seeing that. And I'm curious if other organizations have done any inquiry into this. We're actually just at Kaiser Permanente starting to do some comparisons of this, and I don't have the results of that yet but I'd be curious if others around here in terms of the use are experiencing additional power from the CTM-3.

DR. COLEMAN: Karen, I do have the document. It's from March 14th, 2006.

DR. PACE: Why don't you just tell us.

DR. COLEMAN: Yes, I'm skimming

through to find a better answer to the question.

DR. PACE: Right. Bill came to the microphone so he can tell us quickly.

DR. COLEMAN: Thank you, Bill.

MR. LEHRMAN: In the supplemental information we submitted a week or two ago, the top box correlation between the discharge composite and care transition is 0.29. So again it's a moderately positive correlation, but a lot of difference between, we think, what those two sets of items measure.

DR. PACE: Thank you.

CO-CHAIR MERLINO: Any other questions? Eric, were you going to say something else?

DR. COLEMAN: No, I think Bill's answer's better than what I had teed up.

CO-CHAIR MERLINO: He works for the government. Anybody else? Questions? Should we vote on validity? All right.

MS. ALLEN: Now we're voting on

validity. One high, two moderate, three low, four insufficient. Voting starts now.

Twelve high, three moderate, one low, zero insufficient.

CO-CHAIR MERLINO: All right, feasibility. Actually I see feasibility and usability almost connected because this is a plug and play for HCAHPS with maybe the exception of Estee's comment about is it duplicative for the discharge question. But I mean this could be bolted on the back of the survey and I think pretty much used in the same fashion.

DR. PACE: Right. And I think the other thing is that for organizations not using HCAHPS they can use this alone. And the correlation we saw that Bill just shared with us, the 0.29, would indicate they're related but not duplicative, that they're not redundant.

So feasibility? Any comments about feasibility issues? Okay.

CO-CHAIR MERLINO: Ready to vote.

MS. ALLEN: We're voting on feasibility. One high, two moderate, three low, four insufficient. Voting starts now. All votes are in. Sixteen high, zero moderate, zero low, zero insufficient.

CO-CHAIR MERLINO: Usability and use. Does anybody think this won't be used? Comments? Questions? Anybody from the committee? Shall we vote?

MS. ALLEN: We're voting on usability and use. One high, two moderate, three low, four insufficient information. Voting starts now. We're still waiting on a vote. All votes are in. Fifteen high, one moderate, zero low, zero insufficient information.

CO-CHAIR MERLINO: Overall suitability. Any overall comments, questions, concerns? Vote.

MS. ALLEN: Overall suitability for endorsement of Measure 0228 3-Item Care

Transition Measure, one yes, two no. Voting starts now. All votes are in. Results, 16 yes, zero no.

CO-CHAIR MERLINO: Thank you, Eric.

DR. COLEMAN: Thanks, everyone. Thanks for inviting us.

CO-CHAIR MERLINO: Public comment.

MS. DORIAN: Operator, would you mind opening the line for public comment please.

OPERATOR: Yes, ma'am. If you'd like to make a comment, please press star then the number 1 on your telephone keypad. Okay, and at this time there are no public comments.

MS. DORIAN: Okay, thank you. And are there any public comments from the audience? All right, so we're early. We're ahead of schedule. Did you want to use some of the time to do discussions about --

DR. PACE: Right. We were thinking not a good idea to start a new

measure at this late time, and plus our developers are not expecting us until tomorrow to get into the other measures.

But not knowing how our time will go tomorrow, would you be willing to have a little bit of the discussion about feedback and suggestions based on your experience today? And then if we have time tomorrow we'll certainly continue that.

But maybe just some of your observations, you know, some of the things I teed up this morning about multiple measures in one submission. If you have thoughts about what worked and what didn't work, you know, the process of evaluating measures, what was useful, not useful.

So we're certainly open to suggestions, and we'll maybe just take a few minutes now, because who knows how our time will go tomorrow. And I'll let Lee and Jim start us off.

CO-CHAIR PARTRIDGE: Yes, I will.

I think for a number of us this is new, and Chris, for everybody the PRO-PM process is new and I realize it's very new for the developers. I think from the perspective of someone who's confronting all this information, and trying to make some assessment as a member of the committee, it would be useful if we could be tougher on the developers at the outset.

If when the staff looks at it you don't see the answer in the obvious place, ask the developer to identify it for you, and then we won't all have to go through hyperlinks in trying to find it. I don't know whether that's feasible, but it's almost, if we can't find it we're not going to consider it.

And I don't want to be that mean because I know it's hard from the developer perspective putting this whole package together. So it would be helpful, I think, from my perspective as a reviewer.

MEMBER DOWDING: I would like to

see some sort of two-page summary which basically says we are seeking approval for these measures which comprise of X items. This is the individual patient-level reliability and validity. This is the organizational performance level reliability and validity. This is how many patients we've tested it on. This is the disparity information, if there is any, and this is where it's currently being used.

Just as a summary, just so that we're not having to try and seek through, like you've got the alternative, if somebody's given you the kitchen sink and you're trying to figure what the reliability data is you're supposed to be looking at, or you don't have what you need.

And I think some sort of very short two-page summary. They can give us the calculations behind it, but just the information we need, just as an overview, would be very helpful.

MEMBER LOEB: I would absolutely agree. It's almost the equivalent of an abstract when you're submitting a paper. There's an abstract first that kind of gives you a summary of it. Because for me I mean this is totally new, totally foreign, and hunting through everything and the fact what you said, Lee, of not having it there and then going back and waiting for them to submit it.

I mean when you go and submit a paper you have, you know, one chance. And it is hard. You don't want to be nasty, but it needs to be there. And that's a great suggestion of having some type of review. And they've got all the information in there, so I don't think that would be very difficult at all for them to have just a summary analysis for us to take a look at.

MEMBER CROSS: I think it would be helpful if we could have, as an appendix, just the survey tool or the questions all in one place. Because a lot of the things that we've

read they reference Question 4 or Question 7, and I don't have one easy to place to go find those questions.

MEMBER STILLE: Actually one of my three questions was just asked, so that's good. Yes, I think having the instrument right up front would be good. Two other things real quick. One is just have all the stuff about reliability and validity in one place, because not only is it not apparent, you have to go looking in two or three different places for it sometimes.

And then the third thing, it would be really nice when we come to meet at the meeting, as much as possible, to not have to give any of the ratings insufficient. You know, if there's something that three or four weeks ahead of time the staff says, wow, this data just isn't ready for prime time, to talk to the developers about that so the developers can get it in.

So that when we come here we can

say, yes it's good, no it's not good. Because a lot of the time that we spent today was sort of like, we really don't know if it's good, because we don't know if it's there. So that would be great. Thanks.

MEMBER VAN ZYL: I think the common theme from the developers was, well, they need more specificity. So I think we can hold them to a higher standard of ease with which we find the information that we're supposed to give an opinion on, but not one developer said, tell me exactly what tests you want me to run on what data. And we can't hold them to that if we don't give them that information.

So I think this seems like a new process for a lot of people, but that level of specificity to the developers, I think, would cut down on a lot of confusion on our staff.

DR. PACE: I'll just make one note about that. We've had various expert committees that have looked at measure

testing, and all of the recommendations are not to be prescriptive, because the measure developers are in the best position to know the best reliability testing for the type of data they have, and the sample they have. And same way with validity.

But it is certainly something that we come up against a lot in terms of, you know, requests, just tell me what to do. So it's something that we have to continue to struggle with. And I appreciate what you're saying.

Just so, you know, the other thing that we've been trying to do is to be available to the developers to answer any specific questions they have, if they get to a certain item in the question, or want to talk about testing in general.

But certainly appreciate the comments and what to do, but that's kind of been the push and pull with more specificity in that the experts tell us, you know, it's

impossible to really envision every kind of situation. Because it's a matter of the combination of data, the type of measure, how it's constructed, the data source, you know, so there's a lot of variables that go into picking a particular test. But certainly we're willing to help developers any way we can to give them suggestions.

But, you know, one of the things that we could do, which is what we've kind of provided to you, is just say, you know, the most common approach we see is this, not saying it's required, but at least that might be more useful as well.

MEMBER VAN ZYL: The second comment I have is related, I think. For many of us, interpreting a Cronbach's alpha is tough since we are several years out from our stats courses.

You know, a two-page glossary of common terms would be really helpful, you know, distinguishing type 1 from type 2 error,

distinguishing reliability from validity, and just a quick definition of those terms, I think, would be very helpful for those of us who are more clinical, and less measure develop-y ourselves. That's a word, I've decided.

CO-CHAIR PARTRIDGE: And Chris, you said you had three questions and I think I cut you off after one. No? All right.

Carol Levine?

MEMBER LEVINE: I agree with everything that's been said, particularly the abstract idea. Just get us, I mean I'm used to reading JAMA and all of the other journals. I can understand fairly quickly what's going on here.

This just, you know, is very hard for me. In addition to all of the other things that were said, and I said this earlier but I want to repeat it, that I would like to encourage the developers to think about the kinds of questions that are going to come up,

and answer them ahead of time, like the questions of diversity and ethnicity, and the questions of why did you pick these inclusion criteria? Because they're going to be asked, so it would be better if they thought about those things ahead of time.

There may be other things as well. I sort of lost the sense of the people in all of this process, and I'd want to know, you know, I think that's what this is about. And so we need to understand that a little bit better. Thank you.

Well, one other idea. Sometimes when you submit, for instance, when you submit a grant proposal, often you will see one page that says, have you done this, done that, done this, where is it? You know, so it's a kind of reminder that these are all the things that are essential to include so you make sure that they're in there somewhere. Doesn't say where they have to be, but they have to be there.

MEMBER BIERNER: So that would be

like a checklist. But I think it would be nice to have in, whether it's in a summary, or in the document, what is their intended audience? We know for some of these that are CMS related that they're mandating them, but for many others, you know, what does the developer envision as the intended users of the instrument?

CO-CHAIR MERLINO: Sharon?

MEMBER CROSS: Can I ask a process question just about something that we did today? So the one that we discussed where we only approved the three measures and not the other three, and those will maybe come back to us, will we have the information or a summary of the discussion that we did have already, so that we have that available to us, or we will be kind of having that whole discussion all over again?

CO-CHAIR PARTRIDGE: There will be a transcript of this meeting. But I think it's a good suggestion and, Lauralei, maybe

you can tee up the relevant segment for us. Because at CSAC we often go back and read those transcripts, and they're long.

MS. DORIAN: Well, the transcript will be summarized and eventually compiled in a report so we'll have that. I'll be sure to bring it up section by section.

CO-CHAIR PARTRIDGE: What Sharon's talking about is likely to come up on our conference call in another, what, four weeks or so. And if you can just pull the discussion, or say if you want to refresh your memory it's pages buh-bum through buh-bum.

CO-CHAIR MERLINO: Liz?

MS. GOLDSTEIN: I was just going to add, as from a perspective of a developer, I think it would easier almost if there was one form that, you know, started from the very beginning, the background, and kind of, you know, all the different pieces and then having them multiple forms.

At least from the developer

perspective, it's extremely confusing what goes where, and referencing other parts. So it almost if you kind of went through each piece, you know, background and rationale for measure, and all of that it would be a little bit easier.

I also think the information you gave today about what, typically, you would put in each section, I know there are meetings for the developers. I found, you know, I was on a number of those calls. I found what you did today was so much clearer than the guidance that we had while we were submitting.

So it wasn't very clear exactly what you were looking for for validity and reliability and all of that. And I know there's lots of controversy, like for reliability, do you even look at it at the patient level, since mostly surveys are reported not at a patient level, at the provider level?

So if it is required to do it at

the patient level, kind of giving the rationale for that so developers understand. And that's, you can't proceed without, whether you agree or not given there's lots of controversy, you need it, no matter what. So I think giving some of that guidance early on would have been extremely helpful. I feel like after today I have a much better understanding of what you were looking for than when we were working on the different submissions.

MEMBER NEUWIRTH: I just want to thank our organizers, because I think logistically this, you know, went really well and we were even finishing early which is a nice treat. I guess I have a couple of comments. One is, I feel like, you know, I'm new to this so I'm learning what's sort of expected. But one of the things that is surprising me is how little time we're talking about use.

And so, as part of pre-work we

could be asked to find out, I know we're representing ourselves, but many of us come from organizations or are involved in other, you know, health care systems and such. We could do more to do some discovery work about how it's being used, what the limits are, from, you know, the experts in our own organizations. And coming from a large organization, I think I could have done some more pre-work. So that's like something that it's good for me to know, and that I will use moving forward.

And I feel like, sort of to your point about the people end of this, you know, it's the people in the organizations that feel really missing here about we spent, I think, like to me it feels like 80-plus percent on reliability and validity, and very little time on use.

So I'd love to see, you know, that's really the goal here is to get these measures to be used, to drive performance

improvement. And I don't know what kind of a process there has been in the past of could be moving forward to devote, you know, adequate energy, resources and time on the use element of this.

And then for me, related to that is, you know, the interconnectedness of these measures. And so I think I mentioned this at the beginning, having some sort of a landscape or thumbnail sketch of what we're going to be doing over the next two years, and the landscape of what exists in patient and family centered care so we can see where the gaps are, so we can see where there's overlap and interconnection, and so then we can also go back to our own organizations, and better educate them about what exists and how to continue to drive use, if they're not using some of these measures.

So I think there's tremendous potential here, and it's been a really eye-opening day for me and I want to thank you for

orchestrating it all so smoothly and I'm looking forward to, you know, continuing to get to know my colleagues here as well.

MEMBER BEVANS: Yes, I wanted to address, I think I'm remembering this correctly, so please correct me if I'm not, a specific question that you posed in the very beginning about the value of asking instrument developers to provide information about the perceived importance. Is that right?

DR. PACE: It was about the specific question for the PRO-PMs about how they involve patients or families in identifying what would be of value and meaningful to them.

MEMBER BEVANS: I just want to respond to that specific inquiry to say that I think that is absolutely key, I think, particularly as the patient and family centered PRO people. You know, we should really be encouraging folks to, you know, evaluate that. It's something that has, for

a long time, for whatever reason, been missing from instrument development, especially from PRO development.

The notion of asking the end users, you know, patient or family members what actually, you know, what does transitions to care mean to you, and prioritizing those different concepts to ensure that whatever we're measuring, okay, it's reliable and valid, but does it mean anything? Is it actually meaningful?

And so I just wanted to kind of respond and give my little soap box about that. I think it's really important, particularly for this committee who is charged with ensuring, I think, that patients' and families' perspectives are integrated into quality measurement, that people are not responding to that is not a good reason to exclude it from the criteria.

I think often people don't know how to respond to that and so, you know, maybe

giving them some ideas of what could be done. In fact, I think we saw a couple examples today of qualitative methods that actually answered that question, but it seems like people didn't really know that the focus groups they had done or the consulting they had done with patient and family advisors actually provides evidence for the importance of the concepts.

DR. PACE: And I would ask you, so the question I have about that is so how much information there would be useful? Because obviously all of these experience with care measures that used some type of focus group, and I think probably most of them said that, you know, we had focus groups and did qualitative interviews, but is that enough for them to say that or what level of detail and how that relates now to what the end product is that's being submitted for endorsement?

And that's what I'm trying to, you know, in order for us to give better guidance

to the developers I'm just curious of what your thoughts are about that.

MEMBER BRADLEY: Well, I guess for those of you that are really true researchers and maybe you can answer this, but it seems some of the measures that are coming up now, the research is so old or so dated, and it seems that patients' and families' expectations have changed over the last ten years.

So if we're basing measures on research that was gathered ten years ago, it's almost like should it be updated somewhat or before we move forward? And is there sort of a rule of thumb of if it's over 20 years old or 15 years old, maybe we shouldn't look at it?

DR. PACE: Well, it's a good question. And, you know, for example, clinical, you know, more clinical measures or clinical process measures, that is one of the reasons that we do endorsement maintenance

because clinical evidence may change over time. And it's a good question for us to kind of also ponder of, you know, whether this patient and family kind of discussions needs to be updated, and is that what we should be considering evidence for --

MEMBER BIERNER: No, I wouldn't want us to try to make them redo everything. I mean you want to kind of have an update of these questions still as relevant, you know, maybe things have changed in the delivery of care in some way that the question that might have been relevant ten years ago isn't as relevant. And then if there's been any expansion of their data collection, you know, additional institutions or hospitals or other facilities that they know of or published, you know, would be useful.

MEMBER NEUWIRTH: I think to Katherine's point, maybe in the abstract a simple question of how were patients and families engaged in developing this measure

with simple responses like, focus groups were conducted on X date with the following domains of inquiry, relevance, you know, question construct and so on.

I mean I think that just knowing what was asked is important too, because it's not enough to ask patients, do these questions make sense, but we need to know are they relevant to them. And also like as someone mentioned, we need to know when these questions were asked, because care for patients on dialysis has changed in the last ten years, and their expectations have changed.

CO-CHAIR MERLINO: Debra?

MEMBER SALIBA: Thank you. Again I want to echo congratulations for managing a lot of complex materials. I'm also going to echo a couple of other comments. I do think having a single integrated report on the measures would be really helpful as a reader of these, as opposed to the multiple different

documents. There seemed to be a lot of repetition, and redundancy and it was a pretty heavy reading burden.

And Carol mentioned earlier in the day, and several other people have mentioned, it would be really helpful to have the actual items to start with. It just sort of grounds you in what it is you're going to be reading about.

And then the abstract approach, I think, would be also useful. And that abstract could include, if it's a revisit or an update of a measure, a specific section of, you know, changes to consider, since last review and that could include some of these. Because some things are still going to stand the test of time.

I mean people don't want to be in pain, for example. There's no reason to sort of retest that pain is an important thing for people at the end of life. So there may be some things.

And then finally, I found that, you know, the first section, there were too many subdivisions within the important section that sort of led to repetition and that maybe they could just give, you know, we could tell them what all needs to be covered in that section but not make them all be in very discrete sections.

Because it ended up being repetitive, I think, in part because as was mentioned by the developers earlier, they weren't even sure what these distinctions were, so they would just sort of give you the same information over and over again in that section. But again I think it was, you know, a lot of material to organize and that you all did a great job trying to get it put together.

CO-CHAIR MERLINO: Sherri?

MEMBER LOEB: I think we're on a great track by forming Person-and-Family Centered. There are thousands of outcome measures out there. And when one of the

conferences I spoke at was with -- I'm totally blanking on it, but he was a big outcomes guy down, and he was talking about outcomes.

And my other half of the speech was about, you know, loving and losing, is the current state of measurement really the answer? And, you know, all the outcomes that are out there that are all disease outcomes, yes, they're important, but they don't address the family.

And the biggest statement, or change that came with my husband was when he said, you know, for 18 years I've been working on standards and outcomes, and patient safety and quality, and I never realized that it never really took into consideration the patient who's diagnosed with cancer, and there's nothing that really relates specifically to that patient and their circle of care, and their engagement and family centered.

So I think this is a great start,

and the more we can get important outcomes that really translate to what's important to the patient, other than the medical outcomes, we need to keep fighting for that.

CO-CHAIR MERLINO: I just want to make a couple of comments, and it goes back to what Carol said earlier of why we're here. We are here for the patient. This is the patient's voice that gets translated in a very objective fashion back to the way we run hospitals and health care systems and home health agencies, and whatever it is we do in health care. So it's very, very important to keep that in mind.

And secondly, just to add on to what everybody has said, it's great work by the staff. It's a lot of information. But also, thank you for all of your hard work. I mean this is a lot of hard work.

And I remember telling Karen and Lauralei when I first got asked to participate, oh my god, I didn't realize how

much time this was going to take. But it is a lot of work, and so thank you for all of your effort.

The third thing is be frank. Somebody brought up about, you know, more usability talk. I agree. But we need to make sure that we don't feel squelched and people have their voice. If you have something to say, if you're concerned we're not giving something enough time, yell out. Let's talk about it, because I think it's very important. The things that we adopt are going to be around for a very long time.

And then finally, for tomorrow, tomorrow, one thing that I would recommend is that when we get to validity and reliability, we really turn it over to you to comment first to drive that based on the staff recommendation so that, you know, we kind frame what the staff as the experts think, and then open it up for discussion.

That would be one recommendation.

And any other recommendations for tomorrow to make the process more efficient and effective?

CO-CHAIR PARTRIDGE: Get a good night's sleep. Have a good dinner. Don't talk about measures over dinner. I want to just echo, Jim said everything that I was thinking, which is very nice, but I want to echo thanks to Karen and to Lauralei and Nadine for all your hard work.

And just to say those of you who haven't had the chance I have had to work with Karen very closely for the last two years probably don't appreciate the tremendous value of the patience with which she took us through all of this today. She will be very much missed.

DR. PACE: Thank you. I would just say that, of course, NQF couldn't do any of its work without all of you, so we really appreciate all of the time and effort you've put into this and your suggestions.

And, you know, we try to have

standardization across all of our measures and projects to the extent possible, but we continue to collect this information from the various committees, and we will make efforts to improve.

You know, sometimes we get conflicting information so, for example, many developers like the fact that we've put some things in attachments. It's much easier for them to work with. So we'll just have to continue to try to balance the feedback and, you know, we're always trying to make things better.

So definitely we value your feedback and we will continue to try to implement some of these things as we get our next chance of relooking at the measure submission form and how it might flow a little better for both the developers and the committee members.

So I'm going to turn it over to Nadine to remind us about dinner for those of

you who can make it. I'm pointing to Lauralei and saying Nadine.

MS. DORIAN: Well, she has the slide, though. But there again we have Mio, 1110 Vermont Avenue, and NQF will reimburse you up to $36 for dinner. So hope to see some of you there even if you can only come for a drink or, you know, a quick bite to eat.

Thank you, and see you all tomorrow. We do start half an hour earlier tomorrow, so we have continental breakfast at 8:00 and then we begin at 8:30. And you can just leave your voting devices by where you're sitting, and the name tags. You can leave everything there, except probably not your computers. But it'll be there tomorrow.

DR. PACE: All right, thank you. And we'll see some of you at 6:00 and the rest of you tomorrow at 8:00.

(Whereupon, the above-entitled matter went off the record at 4:53 p.m.)

F

I

Q

C E R T I F I C A T E

This is to certify that the foregoing transcript

In the matter of: Person- and Family-Centered Care Standing Committee Meeting

Before: NQF

Date: 07-28-14

Place: Washington, DC

was duly recorded and accurately transcribed under my direction; further, that said transcript is a true and accurate record of the proceedings.

Neal R Gross

Court Reporter