NATIONAL QUALITY FORUM

+ + + + +

PALLIATIVE AND END-OF-LIFE CARE STEERING COMMITTEE MEETING

+ + + + +

TUESDAY,
MAY 10, 2016

+ + + + +

The Steering Committee met at the National Quality Forum, 9th Floor Conference Room, 1030 15th Street, N.W., Washington, DC, at 8:30 a.m., R. Sean Morrison and Deborah Waldrop, Co-Chairs, presiding.

PRESENT:

R. SEAN MORRISON, MD, Co-Director, Patty and Jay Baker National Palliative Care Center; Director, National Palliative Care Research Center; Director, Hertzberg Palliative Care Institute, Icahn School of Medicine at Mount Sinai; Co-Chair
DEBORAH WALDROP, PhD, LMSW, ACSW, Professor, University of Buffalo, School of Social Work; Co-Chair
BOB ARCHULETA, MD, Physician, Pediatric Associates*
MARGIE ATKINSON, D. Min, BCC, Director, Pastoral Care, Ethics and Palliative Care, Morton Plant Mease/BayCare Health System
AMY J. BERMAN, BSN, Senior Program Officer, John A. Hartford Foundation
CLEANNE CASS, DO, FAAHPM, FAAFP, Director of Community Care and Education, Hospice of Dayton

MICHELLE CAUGHEY, MD, FACP, Associate Executive Director, The Permanente Medical Group, Kaiser Permanente
GEORGE HANDZO, BCC, CSSBB, Director, Health Services Research and Quality, HealthCare Chaplaincy
ARIF H. KAMAL, MD, MHS, FACP, FAAHPM, Physician Quality and Outcomes Officer, Duke Cancer Institute
ALICE LIND, MPH, BSN, Manager, Grants and Program Development, Health Care Authority*
RUTH MacINTOSH, RN, Continuum of Care Manager, Aetna
ALVIN MOSS, MD, FACP, FAAHPM, Director, Center for Health Ethics and Law; Professor of Medicine, Robert C. Byrd Health Sciences Center of West Virginia University
DOUGLAS NEE, Pharm D., MS, Clinical Pharmacist
LAURA PORTER, MD, Medical Advisor and Senior Patient Advocate, Colon Cancer Alliance
CINDI PURSLEY, RN, CHPN, Administrator, VNA Colorado Hospice and Palliative Care
AMY SANDERS, MD, MS, FAAN, Assistant Professor, Director of Cognitive and Behavioral Neurology, Departmental Quality Officer*
TRACY SCHROEPFER, PhD, MSW, Associate Professor of Social Work, University of Wisconsin, Madison, School of Social Work
LINDA SCHWIMMER, Attorney, Vice President, NJ Health Care Quality Institute
CHRISTINE SEEL RITCHIE, MD, MSPH, Professor of Medicine in Residence, Harris Fishbon Distinguished Professor for Clinical Translational Research in Aging, University of California San Francisco; Jewish Home of San Francisco Center for Research on Aging
ROBERT SIDLOW, MD, MBA, FACP, Division Head, Survivorship and Supportive Care, Memorial Sloan Kettering Cancer Center

KARL STEINBERG, MD, CMD, Medical Director, Kindred Village Square Transitional Care and Rehabilitation Center; Life Care Center of Vista; Carlsbad by the Sea Care Center; Hospice by the Sea
PAUL E. TATUM, MD, MSPH, CMD, FAAHPM, AGSF, Associate Professor of Clinical Family and Community Medicine, University of Missouri-Columbia School of Medicine
GREGG VANDEKIEFT, MD, MA, Medical Director for Palliative Care, Providence Health & Services
DEBRA WIEGAND, PhD, MBE, RN, CHPN, CCRN, FAHA, FPCN, FAAN, Associate Professor with Tenure, The University of Maryland School of Nursing

NQF STAFF:

ANN HAMMERSMITH, JD, General Counsel
KAREN JOHNSON, MS, Senior Director
ELISA MUNTHALI, MPH, Vice President, Quality Measurement
RACHEL ROILAND, Senior Project Manager
JEAN-LUC TILLY, Project Analyst
MARCIA WILSON, PhD, MBA, Senior Vice President, Quality Measurement

ALSO PRESENT:

KATHERINE AST, AAHPM
LAURA HANSON, UNC Chapel Hill
KARL LORENZ, RAND Corporation*
STACIE SINCLAIR, Center to Advance Palliative Care
CARL SCHEFFEY, NHPCO
CAROL SPENCE, NHPCO
NEIL WENGER, UCLA Health*

*** present by teleconference**

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

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

8:36 a.m.

MS. ROILAND: Good morning, everyone, if we could just take our seats, and we'll get started.

MEMBER ARCHULETA: All right. Yes, hi, Bob Archuleta.

MS. ROILAND: Hi, Bob. Welcome.

MEMBER ARCHULETA: Thank you.

MS. ROILAND: Alice, are you also on the phone?

MEMBER ARCHULETA: Yes.

MEMBER LIND: Yes, I am.

MS. ROILAND: I'm sorry, who is that?

MEMBER LIND: Yes, this is Alice. I'm on.

MS. ROILAND: Oh, all right. Thank you, Alice. All right. Welcome, everybody, and thank you for coming and traveling all this way to Washington, DC, today for our Palliative and End-of-Life Care Committee Meeting.

So, and a lot of you have spoken with

me on the phone or through email or with lots of people on the project team, so I'm very excited to meet you all in person, so thank you.

So like I said, my name's Rachel Roiland and I'm a senior project manager here at NQF, and I'm just sort of going to kick us off today, and so I wanted to start with a few housekeeping things.

What you'll see on the -- in your microphones in front of you we have a bunch of handouts that we put out on your desk this morning, and within those handouts we have the agenda for today's meeting and within that agenda you'll see just sort of the order of measures that we'll be reviewing today.

In addition to the agenda, we have printed out the algorithms that we use for several of our evaluation criteria, so if you have any questions about how we arrive at whatever decision we end up arriving at, we will be guided by these algorithms. So if you have any questions please refer to those.

And also the last sheet of paper that we have for you is just the recusal sheet that lists out who might need to be recused for a given measure given that they've worked on it in some capacity in the past. We try to be as open and clear about this as possible, so that's why we have this sheet.

And so a couple of other housekeeping things, we just have the restrooms located just near the elevators that you got off of. If you just take a right after those elevators that's where the restrooms are located.

We do try to be kind and incorporate some breaks into our schedule, and so we have three of those scheduled for today. One is at, our first one is a 15-minute break at 10:30, and then we have lunch at about 11 -- 12:45, I'm sorry -- teased you there. And then that will be provided here at NQF and we'll have about a half an hour for lunch, and then we'll have an afternoon break of about 15 minutes around 3:00.

We do have reservations for tonight

for dinner at 5:45, so a true Midwestern time for dinner for lots of you East Coasters who like to eat at 9:00m but it's very close. It's about a block and a half or two blocks away at P.J. Clarke's, so we have a reservation for 23 so we'd be happy for all of you to join us over there, and we'll be walking over there together at the end of the day.

Just some sort of connectivity issues for those of you with laptops or cell phones that have Wi-Fi access, the username and password is listed up there for the Wi-Fi network.

And we do know that you are all busy folks with lives outside of this room and probably have your cell phones with you, and we just ask that you mute those during the meeting because they can just be a bit disruptive if they go off when we're doing discussion or voting.

And with that -- oh, a few other housekeeping things. These microphones, you really need to get up close to them in order for everyone on the phone to hear you as well as

everyone in the back of the room. So please make sure you speak directly into the microphone.

Also, only about two to three of these -- three of them can be on at the same time. So if we have more than that the other ones won't work. And when you get into discussion you'll see how easy it is to forget to turn your mic off when you're done speaking, so just try to be cognizant of that.

And the last bit of housekeeping is that we have the name tents that we ask if you just could sort of angle them towards us so we can better see who's who that would be really helpful. And then when you want to speak I just would like for you to sort of tilt your name tag up like this, so it signals to us that you want to speak and it'll help the co-chairs just facilitate the discussion a little bit better.

And if you just want to go to the next slide, Jean-Luc. And then just -- I should have done this in the beginning, I'm sorry, but just to sort of introduce all the staff here today

that's here to help you and just guide you through this process.

Again I'm Rachel Roiland. I have Jean-Luc Tilly at -- to my right. He's the project analyst and he's the voting guru. So he will guide us through the lovely voting process that we have set up for you today.

Down at the far end of the table we have Marcia, Marcia Wilson, excuse me, a senior vice president here in the Quality Measurement Department at NQF, so she is our quality measurement guru, as well as we have Karen Johnson our senior director for this project as well sitting right next to Marcia.

And she is the one who has led us through all our workgroup calls and will be our methods guru for today as well. And then over here to my right at the side table is Elisa Munthali, another vice president here in Quality Measurement.

So, all right. And so with that I will turn it over to our co-chairs to do their

own introductions.

CHAIR MORRISON: So good morning, everybody, and welcome to NQF. This is, I think, for all of us a very exciting time as a field. Palliative care has evolved to the point where NQF now has a standing committee on measures and this is the second time that this committee has actually met. In the past, I guess it was five years ago, Karen, or more, this was a temporary committee, so kudos to NQF and thank you for doing that.

My name is Sean Morrison. I am from the Mount Sinai School of Medicine in New York City, and my partner in crime here, Deborah Waldrop, and I will be co-chairing today's session and tomorrow.

And for those of you who have not been on an NQF panel, and I have, let me just run through a couple of helpful hints for the next two days before I have Deb introduce herself. Is that right? Okay.

So our role, quite honestly, is to

make sure things move smoothly and relatively quickly. We will be here until we get through the agenda for today and tomorrow, so that whether we finish early or whether we finish on Thursday morning in the wee hours really depends upon all of you. It is -- we can't stop and we will break today at our regular time, but how late we go tomorrow really depends on the work of this group.

So a couple of perhaps helpful hints, we will ask the lead discussant to walk us through the various components of the measures and each of the measures has two or three additional discussants.

It is not necessary to repeat everything that the lead discussant said if you agree with it. You can simply say concur or point out areas of differences.

Similarly, we are going to have, I suspect, a lot of discussion around the various measures, and believe it or not those discussions can get and feel very personal. It feels very

strange to be looking at a series of quality metrics and have a very strong emotional attachment to them. Believe me, it happens.

So what I would ask and what we would really hope is the following. This is not really a debate where you're attempting to convince the other side about the goodness of your heart and the goodness of the measure, so this is an opportunity for all of us to weigh the various aspects of each individual measure.

If somebody articulates a point, feel free not to raise your tent card and say I agree with this person and I particularly want to highlight how important that point is and how strongly I feel about it -- not necessary.

If you have an additional area of concern please raise it. If you have an additional area of strength feel free to raise it, but Deborah and I are going to try really hard to keep comments really focused and novel rather than repetitive this is really what I believe.

Trust me, if we go the former route we will be here for a very, very long time. And please, please don't take it personally if either of us interrupt and say that's a great point but I think we need to move on. Does that work for people?

The last thing, Rachel was very nice about it, we are all incredibly busy people. Everybody in this room will be multitasking. I understand that. Deborah understands that. Karen understands that. Rachel understands that.

That being said, we do have 15-minute break schedules which are either for the bathroom or for emails that don't have to be answered right away.

So I guess what we would ask is if you take a look at your email, you take a look at your texts, and there's something that really can wait until the break, we'd love it if you could just pile those up and answer them at that time, or even better at lunch. If people are focused the discussions tend to go a lot faster and

smoother.

And the last part is, I suspect a lot of you here around the room, including me, are thinking about what is this day going to be like and what am I supposed to be doing. Yes, I have absolutely no idea either. That's why Rachel and Karen are on either side of us.

And they will weigh in when we get lost and they are both close enough that they can step on both Deborah and my toes simultaneously if things are not going smoothly. So relax, it's going to be a lot of fun.

Deborah.

CHAIR WALDROP: It's really an honor and a privilege to be here in this room among all of you with all the expertise that's surrounding us. Thank you for your shared commitment to making the end-of-life care that we provide in this country better.

I also want to say thank you to the staff. You've been incredibly responsive. I know to me as a new person in this process and

probably to those of you who are also new around the table, they answer emails immediately quickly, or amazingly quickly, and have provided us with great information that helps, really, I think, smooth this process out.

So I too will be depending on them to help us know exactly how to keep you and all of us on track. So thank you for being here. I look forward to the next two days and all that we're going to be able to accomplish together, and I think I turn it over to Ann.

MS. HAMMERSMITH: Thank you. Good morning, everyone. I'm Ann Hammersmith. I'm NQF's general counsel, and I'm going to lead you through the disclosures of interest.

If you recall, we sent you a form in which we asked you detailed information about your professional activities. What we're going to do today, and those of you who have been on the committee previously have done this before, is we are going to go around the table and have you orally disclose any interests that you have

that are relevant to the work before the committee.

Just because you disclose does not mean that you have a conflict of interest. Part of the reason we do this is so that we are open and transparent about the activities of our committee members, open to the public, so that the public is aware of where you are coming from.

So what we are looking for you to disclose is if you have any activities that are relevant to the subject matter before the committee, only if relevant to the subject matter before the committee. We're especially interested in your disclosure of relevant speaking engagements, consulting, research, grants. Again, just because you disclose does not mean that you have a conflict of interest.

I also want to remind you that you sit on the committee as an individual. You don't represent your employer's interests, you don't represent the interests of anyone who may have nominated you to serve on this committee. You're

here as an individual expert.

So with that I'll ask you to go around the table, tell us who you are, who you're with, and if you have anything you would like to disclose. And I'll start with the co-chairs.

CHAIR WALDROP: I'm Deborah Waldrop from the University of Buffalo School of Social Work, and I have no disclosures to make.

CHAIR MORRISON: Sean Morrison from the Mount Sinai School of Medicine. Over a decade ago I sat on the technical advisory panel for the PEACE measures. It was so long ago that I don't remember it, but Rachel pointed it out to me. Otherwise I have nothing else to disclose.

MEMBER CASS: Thank you. I am Cleanne Cass. I am a hospice and palliative care physician. I've been doing this quite a while. I'm currently with Ohio's Hospice as their director for home care services. So we take care of about 300 patients a day at home, about a thousand on our census, and what I do is go out and see them in their homes.

I'm also the chair for the palliative care committee at the American Osteopathic Association, so I do some research allocation funding. Would probably be the only area I can think of that I would have a definite -- a direct kind of possible -- I don't think it's a conflict by any means but we do develop research projects. So thank you.

MEMBER WIEGAND: Great. Good morning, I'm Debra Wiegand. I teach at the University of Maryland School of Nursing, and I practice at Thomas Jefferson University Hospital in Philadelphia. I'm on the board of directors for HPNA, the nursing organization, and I have no conflicts.

MEMBER CAUGHEY: I'm Michelle Caughey. I'm a physician and work in northern California, Kaiser Permanente. I oversee the hospice and palliative care programs as well as hospital operations and some other things, ethics. I have no research or other conflicts.

MEMBER ATKINSON: Good morning. I'm

Margie Atkinson. I am the director for Pastoral Care, Ethics and Palliative Care for Morton Plant Mease Health Care, part of BayCare Health System in the Tampa Bay, Florida area.

I also serve on the National Coalition for Hospice and Palliative Care as president of the Association of Professional Chaplains, and I have nothing to disclose.

MEMBER NEE: Good morning. My name's Douglas Nee. I'm an independent consultant pharmacist out of San Diego. I have nothing to disclose.

MEMBER HANDZO: I'm George Handzo. I'm the director of health services, research, and quality at the HealthCare Chaplaincy Network. I'm a little jet lagged because I actually live in L.A.

And part of my job is to advocate for quality measures and quality indicators in spiritual care and chaplaincy care, but I have no particular involvement or interest in any measures before this group.

MEMBER SIDLOW: My name is Rob Sidlow. I'm the head of Survivorship and Supportive Care at Memorial Sloan-Kettering. I do some part time consultancy work for a national for-profit hospice around their educational content, aside from that no conflicts.

MEMBER KAMAL: Good morning, I'm Arif Kamal. I'm a medical oncologist and palliative care physician at Duke University.

Regarding disclosures, I'm the incoming chair of the quality care committee for ASCO and on the palliative care measures panel task force, both of which are positions I've taken since the measures were submitted for maintenance.

Additionally, I have an AHRQ K grant looking at benchmarks and performance for existing palliative care quality measures which are limited to the NQF-endorsed set from prior.

MEMBER PORTER: Hi, I am Laura Porter and I'm a physician and a Stage 4 colon cancer survivor. And I'm with the Colon Cancer

Alliance. I have nothing to disclose.

MEMBER MACINTOSH: Good morning. I'm Ruth MacIntosh. My background is nurse. I'm a continuum of care manager for Aetna's palliative and end-of-life program. I have nothing to disclose.

MEMBER PURSLEY: Cindi Pursley. I'm the administrator of the Colorado Visiting Nurse Hospice and Palliative Care, and I have nothing to disclose.

MEMBER SCHROEPFER: Tracy Schroepfer, I'm a professor at the University of Wisconsin-Madison School of Social Work, and I have nothing to disclose.

MEMBER RITCHIE: Good morning. Christine Ritchie, I'm at the University of California San Francisco where I'm the professor of medicine there. And I'm the past president of the American Academy of Hospice and Palliative Medicine and currently chair the quality committee.

And my major point of disclosure,

which is actually listed separately, is I served on the advisory group for advanced illness initiative for California's 2020 MediCal demonstration waiver. And we chose a couple of the measures that actually are currently being discussed, and so I'll recuse myself from that.

MEMBER TATUM: Paul Tatum, University of Missouri. I'm a geriatrician with the American Geriatrics Society and a palliative medicine physician.

I don't believe I have any direct conflicts that would influence measures. I put a number of disclosures including that I sit on the board of the American Academy of Hospice and Palliative Medicine and the Hospice Medical Director Advisory Council for Hospice Compassus' now known as Compassus Hospice and Palliative Care.

MEMBER VANDEKIEFT: Hello, I'm Gregg VandeKieft. I'm a medical director for palliative care for Providence Health and Services, a five-state system on the West Coast,

and also serve as the advocacy and awareness strategic coordinating committee chair for the American Academy of Hospice and Palliative Medicine, practice both inpatient and outpatient palliative medicine.

I come from a family medicine background. I have no relevant disclosures.

MEMBER STEINBERG: I am Karl Steinberg. I just had cataract surgery a few days ago. I don't usually wear this kind of stuff. I know I'm too young, thank you. You don't have to say it. All right.

So I'm also outside of San Diego, originally a family physician. I'm kind of a long-term care geriatrician now, SNFologist I like to say because I feel SNFist lacks gravity. But I'm also a hospice medical director for a small for-profit hospice in San Diego.

As far as disclosures, I chair the Coalition for Compassionate Care of California. I don't get any pay from them, but they do have an advanced care planning consulting service. So

I don't know that that would be relevant, but just to put it out there.

MEMBER MOSS: Hi, I'm Woody Moss. I'm a professor of medicine at West Virginia University School of Medicine. I'm a nephrologist and a palliative care physician.

I chair the Coalition for Supportive Care of Kidney Patients, and I also have funding from the state of West Virginia to direct the West Virginia Center for End-of-Life Care.

MS. HAMMERSMITH: Okay, thank you. I understand that there's some committee members on the phone, so I'll call your names.

Amy Sanders. Is Amy Sanders on the phone?

MEMBER SANDERS: Yes, I'm here. Sorry. My name's Amy Sanders. I am a geriatric neurologist at SUNY Upstate Medical University in Syracuse, New York.

And I have no direct disclosures, but sort of indirectly I'm a member of the American Academy of Neurology Quality and Safety

Subcommittee which has developed measures, none yet directly related to palliative care that are standalone, although there is a measure about suggesting advanced care planning for dementia patients that is part of a measurement that is currently in development.

And I obviously participate in advanced care planning discussions with my own patients, but that's it.

MS. HAMMERSMITH: Okay, thank you. Alice Lind.

MEMBER LIND: Hi. This is Alice Lind. Sorry about my terrible voice. I'm at the Washington Medicaid agency called the Health Care Authority, and I have no disclosures to report.

MS. HAMMERSMITH: Okay, thank you. Bob Archuleta.

MEMBER ARCHULETA: Yes, Bob Archuleta. I'm a medical director of Noah's Children Hospice and Palliative Care, have been for 19 years. It's now a program of the Bon Secours Catholic health care system, and I have nothing to

disclose.

MS. HAMMERSMITH: Thank you. Eduardo Bruera.

(No response.)

MS. HAMMERSMITH: Okay, all right. Thank you, everyone, for those disclosures. Do you have any comments or questions of each other based on the disclosures this morning?

Okay. Before I leave you, I just want to remind you that in order to make our conflict of interest process work, we rely on all of you as committee members.

So if during the discussion you think that you may have a conflict of interest, you think that a committee member has a conflict of interest or is behaving in a biased manner, we ask you to speak up in real time. You can do that in the meeting if you wish.

If you don't want to do that you can approach your chairs who will work with NQF staff or you can approach NQF staff directly. And I also just want to say that the advice that Sean

gave you was dead on about how to discuss things within the meeting.

So thank you very much, and have a good meeting.

MS. JOHNSON: So good morning, everybody. Thank you again for coming. It's my pleasure to go over our portfolio with you, and just remind you of a few things in our evaluation process before we really delve in and get into the measures.

Sean stole some of my thunder though, so that's okay, just wanted you to know. Let's go to the next slide, please.

So we want to do a portfolio review basically to make sure that you understand what we have in our portfolio. And when we talk about our portfolio, we talk about all of the measures that are endorsed by NQF in a particular area.

You guys are only seeing some of them in this particular project and that may have to do with timing or what-have-you, but for this project we're looking at measures related to

palliative and end-of-life care.

Our measures that we endorse should be useful for accountability and public reporting purposes for any population in any setting of care. We will be looking at assessment in management, physical, psychological and spiritual aspects of care. We have a couple of care planning measures that we'll be looking at, appropriateness of care measures, and the new ones are the experience-of-care measures from the CAHPS survey.

We currently have more than 36 endorsed measures within the area of palliative and end-of-life care, and you'll see in a few minutes why I had to say more than 36 rather than give you an actual number. Let's go to the next slide.

So as Sean as mentioned, this is not our first foray into this topic area. Ten years ago now we actually pulled together a group, and I'm not sure if anybody was on that original group. Was anybody on the 2006 group? Doug was,

okay, great. Welcome back, Doug.

As part of that -- and Cleanne, okay, and Sean. Oh my. See I should remember these things, and I apologize that I don't.

So in that work the group came up with a measurement framework, and we'll talk about that a little bit more later. It was a very comprehensive framework. It discussed scope, specifically recipients of care, the various care settings, the professionals who provide that care.

It talked about structural elements of care, for example, having team-based care, having educational programs, et cetera. It articulated the domains of care, I believe from the National Consensus Project domains. It talked about levels of measurement and outcomes.

So again, a very comprehensive framework, and along with the framework that committee also came up with what we call preferred practices, 38 preferred practices that were organized by the domains and the IOM

dimensions of care, and those practices are actually still NQF-endorsed.

We endorsed those practices, so maybe one of the things that we'll talk about later today or probably more tomorrow is what do we need to do with those preferred practices? Have they kind of done their job and they can go? Do we need to keep them? Do we need to revisit them? So that'll be some of the work that we'll talk about tomorrow.

In 2012, we pulled together what we called a steering committee at the time. It was a temporary committee, and during that project 14 measures were endorsed.

And you're seeing, I think, most of those come back as what we call maintenance measures this time around, so having another look at them four or five years later to see if they still meet NQF's criteria for endorsement.

Finally, we have a measure applications partnership, PAC/LTC workgroup. Let's see if I can remember what PAC/LTC is.

It's post-acute care/long-term care workgroup, and they give advice and feedback to HHS on programs for post-acute care and specifically hospice.

So in 2012 that group actually wrote a paper. Had a meeting and came up with some strategy and they noted several gaps in care. They noted what we had at the time, and that was pretty much going on at the same time that the endorsement project was going on.

So back then there wasn't a tremendous lot of measures, and that comes through in that report. But we have MAP meetings every year, so we get ongoing feedback from the MAP about these measures.

And let me see if I can remember who's on the MAP. I know Sean is on the MAP and he is on the hospital workgroup. And Alice, I believe, you're on the PAC/LTC or are you on duals?

MEMBER LIND: I am on duals.

MS. JOHNSON: You're on duals, okay. All right, so we may not have anybody on the MAP

workgroup.

You'll notice in some of the preliminary analyses that we've provided you, in the use and usability section, we talked about MAP feedback. So this is part of what the MAP feedback is. If you would like to know more about MAP and what it does and that sort of thing, we're happy to explain that to you. We did a little bit of that in your orientation.

It's probably not critical that you know that, but it is feedback from folks similar to you who are interested in measures for particular programs run by CMS. Go to the next slide.

So this is a diagram that I'm sure everybody's familiar with. We see palliative and end-of-life care as being on a continuum alongside, often, disease modifying treatment with the idea that palliative care can be provided at any time during a condition or illness, but often it steps up as the disease progresses.

At some point along that trajectory we start talking about end-of-life care, and we also realize that end-of-life care actually goes beyond the death of the patient and there is grief and bereavement support for family members.

So I will be interested in hearing if anybody feels like that this particular diagram is outdated, outmoded, or is it still good to use as an illustration of what we mean by palliative and end-of-life care. Let's go to the next slide.

So first of all, apologies. You can see that I'm not a graphics designer, and I'm not even sure that you can read this. But the idea here is we wanted to have a measurement framework, so we still have the framework that you guys -- some of you, Cleanne and Doug, Sean, created in 2006.

But that one was extremely comprehensive. There wasn't a picture, and we were wondering if we need to modify it in any way, can we draw it out. This is where we landed

so far. Notice the word draft. We'll be getting your feedback as to whether you think, you know, is this the way we want to go, do we want to do something else, what the picture looks like. We have graphic people here that could draw us a pretty picture.

But the idea here is that the patient and family are at the center of the care. We have kept the NCP domains of care. You can see those, the psychological, physical, cultural, social, et cetera. Those are in the next ring out.

We're recognizing the various settings of care and we have five of them there -- inpatient, outpatient, hospice, home, and SNFs, although I should change that to NFs, sorry. And then the outer environment if you will is end-of-life care, palliative care, bereavement care, so those may or may not be quite the right ones.

The idea of having a measurement framework is to help us understand and categorize or classify the measures that we have and the

measures that we don't have.

So we want to be able to use something like this to help us kind of quickly get our minds and our hands around what is there and what isn't. And that's the idea of the framework. Let's go to the next slide.

So what is in our portfolio? I think I talked about these in our orientation but I'll talk about them again just to bring them back to your mind. Most of the measures that we have are around physical aspects of care, and of those, most of them are dealing with pain.

And all 14 of these -- no, I take that back. Some of the pain measures we'll be looking at in this project, the two dyspnea measures and the -- I called it constipation. It's the bowel regimen measure we will be looking at in this project.

We have a couple of health-related quality of life measures and a couple of cultural aspects of care measures. Those -- the stars indicate that they are not in this project. They

are in other projects, so we won't be spending much time on them other than I wanted you to know that they exist. They are out there.

We have the spiritual, religious and existential aspects of care. We have the one measure right now for that in the hospice and palliative care setting.

We have three measures that I have placed under ethical and legal aspects of care. They are specifically related to care planning. Now one could argue that, you know, I put them in the wrong domains. It's probably not worth an argument. I can move things around if you want me to. I'm not, you know, totally wedded to all of these. Two of the care planning measures are part of your purview today. Let's see the next slide.

Care of the imminently dying patient, we have six of those and they are all what I've noted as appropriateness of care measures. These are the ones, you know, ED use in the last 30 days of life, ICU use, et cetera.

And then we have experience of care measures, and this is where it got tricky. Notice I didn't have an N there. And I did this, really, because I'm considering some of the other measures from the various CAHPS surveys as part of the portfolio even though they aren't specifically related to palliative or end-of-life care, but it could be encompassed in those things.

So we have CAHPS measures for hospital, both adult and child clinician groups, dialysis facilities, home health, and nursing facilities.

So just like the hospice CAHPS measures that are in the project today and tomorrow, each of these measures or instruments, I should say, have several, sometimes 10, 11, sometimes a few fewer, performance measures underneath. And they all are very similar in ways, communication, et cetera, respect, things like that. So I have put them in our portfolio for now.

We also have two other experience of care measures, the bereaved family survey and the family evaluation of hospice care. Both of those were evaluated about a year and a half ago in our person and family-centered care projects.

One reason is because they were what we call patient-reported outcome-based performance measures, and we had a project that looked only at those kinds of measures and it had been a while so the timing worked out. We didn't know that we would have a palliative care project at the time, so they were slotted with that one. Okay, let's go to the next slide.

So gaps in our portfolio, so based on the framework, that kind of greenish circle that you saw, it's really easy to see that some of the domains of care are not really represented, particularly the social aspects, and for care of the imminently dying I believe all the measures that we have right now are limited to cancer patients only. So that is a gap that, you know, we probably may need to fill.

We do not have measures specific to bereavement care or measures that are applicable specifically to the family or caregiver. So the CAHPS measures ask the caregiver about the care that was given to the patient, but there's nothing yet that actually measures the caregiver themselves or their experiences.

From the MAP work articulated in the 2012 work, and to some extent repeated more recently, they noted gaps in terms of access to care particularly palliative care. They talked about need for goals of care and shared decision-making, and also we have no costs of care measures.

So those are at least some of the gaps. Tomorrow we'll be talking a lot more about gaps and about strategies for how we want to start thinking about filling some of those measurement gaps. Next slide, please.

So Sean has already -- this is some of the thunder that he stole so I don't really have to go over this too much. Some of your

responsibilities as standing committees is to provide input on the relative framework or frameworks. You may know of other frameworks that you think would do better, so we'll be talking about that tomorrow as part of our discussion, and I really do want to hear what you think about that.

The other thing that I wanted to highlight from here is you really are overseeing our portfolio of palliative care and end-of-life care measures, so with that we want you to think about the portfolio as a whole.

You're going to be evaluating the individual measures, but keep them in context of what's out there, okay. That's what we want you to do. We want you to know what's going on in the topic area so I've shared some of those things with you, and really we want you to think about and provide us feedback again -- mostly tomorrow -- about how we want the portfolio to evolve, okay. Let's go to the next slide.

So this is a bit of an example. So

remember I said that we had 11 pain measures in the portfolio. What you see here is ten of those. The 11th one is kind of a different animal so I didn't include it here, but what you see here is we have several different process measures as well as a couple of outcome measures. We have them across settings.

So part of the question when you're thinking about the portfolio as a whole is, you know, once there's outcome measures do we still need process measures? We're not saying that we don't, but we're saying think about it, okay.

Hospice, hospital, home health, et cetera, for the home health and the nursing facility settings those are constrained to some extent by the assessment instruments, the MDS and the OASIS, but the other ones maybe aren't so much.

So do you really need screening measures, for example, different screening measures in the different settings? So again another thing to think about, maybe you do, but

these are some of the conversations that we want you to think about and that we'll have later today and possibly tomorrow. Next slide.

So overview of the evaluation process, I just want to remind you of a few things. Let's go to the next slide. Your role is as a proxy for our membership, working with us to achieve the goals of the project, and for today that means getting through all of our measures as Sean has mentioned before.

You will evaluate each measure against each of the criterion and you'll rate the measures. And I believe everybody was able to attend the workgroup calls so you had at least a chance to practice and to think about those things. Today you'll actually be voting, and Jean-Luc and Rachel will take us through that. You have your clickers.

You'll be making recommendations to NQF membership about suitability for endorsement. That's what your job today is to do. As I already mentioned, you will be the overseers of

our portfolio of these measures.

Finally, select two-year or three-year terms. So this is what we call a standing committee, what that means is we want to have continuity over time. So we don't want you guys to go away after tomorrow and you're done with us. We want you to hang around for a while.

So what we're going to do probably tomorrow is we're going to pass the hat, and we're going to let you reach in and choose a two- or a three-year term, okay, so it'll be random. If you don't want to have a three-year term, let us know if you'd rather choose the smaller term. Once your term is over there is a possibility to re-up for another term as well, so just FYI. That'll be a little fun tomorrow. Okay, next slide.

We have a few ground rules for today's meeting. I think there's really nothing here that is surprising. I think Sean has hit all of these things. I'll just pause for a second and see if anybody has any questions about ground

rules or anything like that. Do these look reasonable to everybody? Okay. Next slide, please.

So measure discussion and voting, the way that we will be running this portion of the meeting is for all of the measures -- and forgive me, I'm going to go to my notes here to make sure I cover everything.

We will start each measure by giving our developers a chance to introduce the measures. We are fortunate that our developers come and interact with us at our various meetings, and you'll notice that we've actually provided a place for them at the table. So they will sit here when they're discussing their measures.

And we've done this so that they can more easily respond to questions from you guys, okay. So just like you guys they have their little tent card and they can put it up if they have something that they want to ask or input, and our co-chairs who are facilitating will

select them to speak when it fits them.

Sean's already mentioned the lead discussants. You'll be discussing things as we did on the workgroup calls for each of the criteria. So we'll ask you to do a brief summary of the pre-meeting evaluation comments and/or workgroup discussion, because sometimes the workgroup discussion went off into other things, right. So any of those kind of things mainly emphasizing areas of concern.

And what we'll ask you to do and we'll work through it in our first measure, but we'll have you start with evidence and talk about evidence and then stop, and then we'll have discussion, okay. And then when it's time to go back and we've voted or done whatever we needed to do, then we'll go back and somebody can talk about gap. So we'll proceed in the direction of our criteria.

You'll recall that we have, as staff, provided preliminary ratings for you. Again those are not binding on you. They were intended

to be used as a guide to facilitate discussion, so we just want to make sure that you understand that it's not binding. And again the lead discussants and the workgroup discussants were the ones who did the really deep dives, but we really want everybody to be able to discuss and contribute to all of the measures. Okay, let's go to the next slide.

Overview on voting, so I've already mentioned this. We'll vote by criterion in the order that's presented. We have some of our criteria are must pass, some are not.

At the end there will be a vote for overall suitability for endorsement. You guys did not do that vote or discuss that on the workgroup calls, so that'll be new from what you did before, okay.

If a measure actually fails on one of the must pass criteria, then we stop discussion and voting. So if something goes down on evidence that's it, we do not keep going, all right. We'll just go to the next measure. Let's

go to the next slide.

Now as a reminder for those of you who are totally new to the process this isn't a new -- everything's new to you, right? But those of you who have been around for a while, we are treating maintenance measures a little differently than we have done in the past and our new process is listed on the slide, less emphasis on evidence if it's unchanged.

Okay. Now that doesn't mean that we won't discuss evidence. There will be an opportunity to discuss it if you want to. You may want to re-vote on evidence and that's fine. You may not want to. That is also fine, okay.

There will be an increased emphasis on gap or opportunity for improvement. We're not making any difference in the way that we're treating our specifications. Specifications are always something we want to look very closely at and make sure we understand and are comfortable with those.

That said, there is less emphasis on

reliability and validity testing unless that testing has not been updated. So if the developers have not updated their testing, they haven't done anything new, there may not be much of a reason to have long, drawn-out discussions or even votes on those. However, if they did update their testing then we will discuss and we will vote. Okay.

Feasibility emphasis is unchanged, but we have increased the emphasis on usability and use. Again this is for our maintenance measures. And just again to repeat. Discussion and/or vote on evidence, reliability and validity may not be needed. Let's go to the next slide -- and don't worry about it. We will walk you through all those. We'll point out when you may want to think about it or not think about it, so we'll help you with that.

A couple of other considerations, I think we touched on this in at least one or two of the workgroup calls. We do have a possible exception to the evidence subcriterion that you

can invoke so that might be something that we need to do.

And endorsement with reserve status, this is a special status. It is an endorsed status but it's pretty much an inactive status for measures that are topped out. That means that everybody's doing so well that there's really not anymore opportunity for improvement.

The caveat with that is that the measures have to pass all of the other criteria, and I'm not sure that reserve status is going to be an option. It might not be something we need to talk about, but I wanted to alert you that it might come up, and if it does, I'll give you more details when we get there. Let's go to the next slide, please.

So voting tools, and in a minute I think Jean-Luc is actually going to do a practice session. I don't know if we want to do it now or if we want to do it closer to time, but you will have a tool. You have a remote clicker. And -- yes, Sean is demonstrating our clicker.

Those who are on the phone are actually going to be communicating with Jean-Luc and Rachel and they will be casting their votes for them. So you'll see Jean-Luc and Rachel clicking. That's because they have their -- we're not padding the votes. We will be taking those from our remote members.

So the instructions are to point towards Jean-Luc. The remote will briefly display your choice and basically you can click that thing as many times as you want. It's going to record and keep the very last vote. I'm correct on that, right, Jean-Luc? Okay, great. Let's see the next slide.

Now let me tell you about consensus before we get to our little fun exercise here. Consensus, what's it mean to pass, right? So first of all we have to have a quorum. We have to have at least 66 percent of our committee, and looking around we do have that so we don't have to worry about that.

It can get a little hairy toward the

end of Day 2, you know, when people have flights to catch and whatever, so we do pay attention to how many people are voting. We want to know that denominator so we can get our percentages.

To be recommended, measures must have greater than 60 percent of the committee yeses. And just to remind you, a yes or a pass is high or moderate together. If it's less than 40 percent passing then something's not recommended.

And that lovely area, 40 to 60, inclusive, is what we're calling consensus not reached, and you'll hear us use the term gray zone, perhaps, if any of the votes fall in that gray zone. Any measures that fall in the gray zone, we will actually move forward. So if we're voting on evidence and the votes land in the gray zone, in this consensus not reached zone, we will continue the discussion and go on to the next criteria and keep going even though it hasn't officially passed, okay.

So our must pass terminology is a little bit maybe not as accurate as it used to be

when we instituted the gray zone. Next slide.

All right, that's it. Jean-Luc, let's go ahead and do our little voting exercise.

MR. TILLY: Sure, as Karen described, just take your clickers, and the polling is now open for what does CDP stand for, four options.

MS. ROILAND: And again you have to point them at Jean-Luc to get the computer to register the vote.

MS. JOHNSON: Hopefully you're seeing your clicker light up. If you're having a problem, let us know. You can see in that circle there, 22 people have voted. I think --

MR. TILLY: Yes, just one time is good. So we have 23 votes, which is what we're looking for. All right, so the results are 2 for committee development program, 17 for consensus development process -- congratulations -- 2 for careful deliberation process, if only, and 2 for consensus development program.

MS. JOHNSON: Okay, thank you, Jean-Luc.

And before we hand it over to Sean to take us through our first measure, I know we had at least one new person come in, so -- hi, Linda. You weren't here in our disclosure process, so can you just introduce yourself and let us know if there's any potential conflicts or things that you'd like to disclose.

MEMBER SCHWIMMER: Okay. Good morning, everyone. I apologize for being late. Linda Schwimmer, I'm with the New Jersey Healthcare Quality Institute. We're a not-for-profit so I think I serve as a consumer role, which is a role, and I've no conflicts to disclose, no conflicts.

MS. JOHNSON: Thank you, Linda. And on the phone, we didn't have Eduardo earlier. We still don't think we have him. Okay.

All right, are you guys ready to jump in? All right, Sean.

CHAIR MORRISON: Terrific. So we are going to begin with a dyspnea screening measure which is 1639. 1638, thank you, from University

of North Carolina at Chapel Hill.

And as Karen said, the way that this will go we'll have the measure developer give us about two to three minutes' overview about the measure, then I will ask the discussants to go through.

Doug. Is that not what I said? Dyspnea treatment, sorry. This is, the last time I did this I didn't need glasses.

Dyspnea treatment, I'll ask the discussants to go through each of the criteria. After each criteria, we will vote and then we will move forward.

Again just a couple of quick reminders, if you have a question, tent card goes up. We will get to everybody, don't worry. When you finish your question tent card down, and what everybody will forget is please turn your microphone on and off.

And finally, the folks, Amy, Alice, and Bob, it is really hard for us to see your tent cards, so please just jump. Don't be shy,

just jump in. I will put you on the list and we will get to your comment. But just as I said, don't be shy. Just say I have a comment and we'll go from there.

Did I miss anything, Rachel?

Dr. Hanson from the University of North Carolina.

MS. HANSON: Thank you all so much. It's a delight to be here after a very early morning flight from North Carolina.

I am going to start, because this is the first of a series of measures that I'm measure steward for, to give a little context of the development of that particular group. So forgive me and I promise I will never talk this long again when introducing measures, but just wanted to give a little context.

So the 1638 dyspnea treatment measure is one of a group of quality measures that were developed in the PEACE project. And the PEACE project was funded by CMS in preparation for quality measurement in the hospice population,

but a very specific charge in the PEACE project was to select quality measures that were relevant to both hospice and palliative care populations.

So when these measures were put forward for NQF endorsement, we utilized initially the data that came from the PEACE project itself in developing and initially testing the measures in the hospice population, and then we augmented that with data from a second deployment of a subset of the PEACE measures.

All of the currently endorsed measures were included in that subset to collect data in a hospital-based palliative care population. So in your materials in which you see the original submission of this group of measures, you see data both from originally a hospice group and separately a palliative care data source.

In this maintenance submission, we have primarily provided data that is new for the hospice population because these measures have now been incorporated as originally intended in

the HIS, in the hospice quality reporting system.

The measures themselves are also being incorporated in a variety of ways in palliative care populations although none of that data is ready for submission at this time. And I'll just give a general comment about that and then when I talk about specific measures, I can give you a little bit more about that.

So in the more proximate future, several of the PEACE measures are being utilized in a CMMI project, CMMI-funded project, an innovation project that is collecting data on a palliative care population in a variety of settings.

There is some data collection moving forward through the PCRC, the Palliative Care Research Cooperative group, on some of these measures again in a diversity of settings with palliative care populations.

And then in the more distant future, these measures, some of these PEACE measures are being incorporated in a QCDR, so in a qualified

clinical data registry which is being submitted to CMS, and this fills a gap in palliative care practice, specifically, where outside of a hospital setting, palliative care providers are struggling to demonstrate quality measures to meet new payment criteria that they must meet. And at the moment they're forced to use quality metrics that are really not relevant to the palliative care population.

And then, finally, these measures are also being incorporated in the evaluation process of the large Medicare demonstration project that is just now being initiated but which will test a model of concurrent hospice organization-delivered palliative care along with disease-directed treatment in a concurrent care model that will actually match a lot of what happens currently in palliative care practice.

That project and data from that project is probably five years hence. But I want you guys to have that landscape just to understand the context, because I know that one

of the initial questions is why are there two data elements that segregate the hospice and palliative care populations.

So the dyspnea treatment quality measure, hopefully you all see this in terms of its specifications, but basically it's a percentage of hospice or palliative care patients who initially screen positive for dyspnea as a clinical problem and who then, the numerator criterion, go on to receive treatment for their dyspnea within 24 hours' time.

This quality measure was submitted to NQF as a paired measure, and you'll hear that from me again, with dyspnea screening. In developing these PEACE measures, we made the assumption that the process included screening for a symptom first, which is what we do by simply asking if it's present and asking if it's a problem for the patient, if it's something that's distressing to them, and then going on to assessment and treatment based on that establishment of a denominator population.

So that's what this measure reflects and it's paired with a dyspnea screening measure that you'll hear about. In this submission, we are asking for a modest change in the specifications of the measure.

So when originally submitted, this measure excluded individuals enrolled in hospice care who were in hospice for seven days or less. That was at the advising of the expert panel that we put together for the PEACE project itself, thinking that perhaps seven days were necessary for hospices to accomplish dyspnea treatment, but there was a lot of concern that excluding that population, as we all know, excludes a very high percentage of people who are enrolled in hospice.

So in this maintenance submission we are asking and have provided data to indicate that that exclusion can be removed so that it can apply to all persons enrolled in hospice regardless of length of stay.

I think I'll stop there. There may be other questions or comments. Just to ask the

chairs, do you wish for me to stay here during the discussion?

CHAIR MORRISON: Yes, afraid so.

MS. HANSON: Glad to do so.

CHAIR MORRISON: Questions on process before we take a deep breath and begin? Indeed not. Okay, so I have Cindi and Ruth as the discussants.

Cindi, do you want to walk us through the evidence piece? We'll start with evidence.

MEMBER PURSLEY: Okay, can you hear me? Okay. And just so you know, I may be from Colorado -- this is a cough drop not chew.

This measure is for patients that screen positive and have treatment initiated within 24 hours of the screening. It's a process measure. It's in for maintenance.

They do have the results from the last -- gosh, I'm going to say almost three years now from the hospice item sets submissions, and so the data is there for this measure.

So the evidence initially and the

evidence that's been collected still supports this measure.

CHAIR MORRISON: Ruth? Where's Ruth? There she is. Anything to add for --

MEMBER MACINTOSH: The evidence remains but it has been updated by the developer with two new guidelines, so there's plenty there.

CHAIR MORRISON: Well, good. And we don't have to vote, right, because -- sorry.

MS. JOHNSON: So the question that we would have for you guys as the committee is do you know any other evidence for this that would be contradictory in nature?

So they provided updated evidence. It pretty much just strengthened the story from before. If you don't know of anything new that would make you change your mind or shed some kind of doubt on this measure, then we could decide not to do any further discussion and not to vote.

CHAIR MORRISON: Anything new, folks? Terrific. Then why don't we move to the next criteria which is the opportunity for

improvement.

And again on Cindi and Ruth, and just a reminder and I will do this just for the beginning, this first measure. This is describing any data on current performance, opportunities for improvement, any data on disparities and any other issues in terms of questions to the committee from the preliminary analysis.

So can we go back to Cindi?

MEMBER PURSLEY: Yes. The -- from a user standpoint an area of an improvement would be the way the question is asked. It's asked if treatment, if the patient screens positive for dyspnea, was treatment initiated within 24 hours.

And what I found is that my staff, who are very literal, if the patient was on treatment and they did not initiate a new treatment, then they answered it as no, and it made it look like the patient screened positive for dyspnea but did not receive treatment.

So my recommendation for an

opportunity would be, for example, with the bowel, the opioid bowel question, it says did you continue or initiate a bowel program for a patient on opioids? Any my suggestion would be to add, did you continue or initiate treatment for dyspnea?

CHAIR MORRISON: And I guess, Ruth, your comments?

MEMBER MACINTOSH: I agree with Cindi.

CHAIR MORRISON: Open for discussion. Karl.

MEMBER STEINBERG: I also agree, but I might say initiate or continue instead of continue or initiate.

CHAIR MORRISON: Okay, others? Laura.

MS. HANSON: I'll just point out, and I do agree that that's language out of the HIS in its implementation, which I think shows for all of us some of the complexity of implementation of quality measurement in real-world practice, the use of the term initiate was very purposeful in indicating that once the hospice provider assumes

responsibility for the care of the individual, they initiate all care. They initiate it.

However, from a provider's perspective, it may well be that one continues a treatment that was begun prior to assuming responsibility for care of the patient.

So it's an implementation issue, I think, rather than a specification issue. It doesn't change the measure specification, but it shows the subtlety of implementation.

CHAIR MORRISON: Arif.

MEMBER KAMAL: So, you know, the concern here is that it's a patient presumably who's been on a stable regimen of something for dyspnea, and despite that has, you know, moderate to severe dyspnea, you know, on admission to wherever they are, hospice or the hospital, which, you know, in terms of expectation-setting may be a four out of ten as to where they're going to be and where the best of our evidence says that they may go.

At the same time they're saying we're

just going to continue on what they had before with no, you know, sense of trying to tweak things or modify things for find things or add things. It is a little troublesome to me, too.

So I think, you know, initiation or a continuation with the spirit that we are really trying to maximize what's possible is really, I think, what we're trying to get at.

So I don't think it can be purely, you know, the continuation part. I think the initiation is actually a really important part of it, because you're saying somebody has a remarkable amount of dyspnea that's being noted.

CHAIR MORRISON: Other comments, thoughts?

So I'm going to turn to Rachel for a second and ask, because at this point for this particular area sometimes we would put forward a measure -- or, sorry, a motion not a measure that would be voted on to change to identify the gap.

And I'm going to ask Rachel, would this be something that we would put forward? Is

this a recommendation?

MS. ROILAND: And I'm actually going to turn it over to Karen to answer that one.

MS. JOHNSON: We can certainly put that in the minutes of our report. And Laura has heard, and it sounds like it's something -- let me make sure I understand. This is something that you can actually change in our specs, or you're saying that it's really, the confusion is more in the HIS instructions?

MS. HANSON: It really is in the HIS instructions, and that's what I mean by implementation. This would not be a change to the specifications of the quality measure itself, but clarifying implementation is a welcome contribution. I just want to make that distinction.

MS. JOHNSON: And we will reflect it that way in the report.

CHAIR MORRISON: Perfect. So I think we are -- do we need to vote on this? There's nothing to vote on, correct?

MS. JOHNSON: Well, you still haven't actually talked about opportunity for improvement.

CHAIR MORRISON: Oh, yes. I apologize. Opportunity -- Cindi.

MEMBER PURSLEY: Would it be better served to say initiated or adjusted because it's then proactive? It shows that you're doing something.

CHAIR MORRISON: Christine.

MEMBER RITCHIE: Just as a question for clarification, when you're talking about opportunity for improvement are you talking about adjustment of the measure itself or the gap that the measure demonstrates?

CHAIR MORRISON: I'm sorry, the gap that the measure demonstrates, not the measure itself. Yes, I think as Rachel told me we have an hour, so we're going to let this conversation go a little longer than I normally would. But yes, we're looking for opportunities in terms of gaps and not to rewrite the specs of the measure,

per se.

Arif.

MEMBER KAMAL: Sean, this is going to come up, and this is to Karen and Rachel, too. In terms of changing specs, words, phrases, et cetera, where or if is that germane within the total agenda of the conversation, because I have a feeling that's going to come up repeatedly.

MS. JOHNSON: Basically, in terms of your voting you should be voting on the measure as it's written. However, if you have recommendations or ideas and it's something relatively simple, you know, a little wordsmithing, and if the developer is amenable to that then they could do that. But you pretty much have to vote on what you have as is.

So we want to make sure that we capture any of your recommendations, but this is not, you know, our opportunity to rewrite measures.

CHAIR MORRISON: Laura.

MS. HANSON: And I just want to point

out that the language in the measure specification is receive treatment. So I want to make that distinction. The language in the specification is receive treatment, which I think captures very much the intent of what your concern is.

But implementation through the HIS uses a somewhat different language because of concern about the passing of responsibility if somebody comes into hospice, they screen positive for dyspnea, meaning they have dyspnea as an active problem on the moment of admission, making sure, and I think this goes a little bit to Arif's comment that treatment is initiated within 24 hours to address that symptom distress, but the specifications for the measure are really to receive treatment within 24 hours of dyspnea being identified as an active problem for the person.

CHAIR MORRISON: So I will jump in and I will not do this again. But I guess, you know, one of the opportunities in gaps is, as Laura

described, in getting the lack of the new palliative care data that we had that hospice data are provided, but both in terms of disparities and how this is performing within the palliative care population because a third of hospices are now above the 90th percentile, the question is whether that is an opportunity or a gap.

Christine.

MEMBER RITCHIE: Just so I understand what you're saying, Sean, you're suggesting that it could be almost considered to be, quote-unquote, topped out in the hospice community, but because we don't have data in the palliative care community there still probably is a lot of opportunity for improvement? Is that what I'm hearing you say?

CHAIR MORRISON: That's what you're hearing me say. And I think we also heard from the developer that those data will be coming in shortly, but they're not available now. So thank you for summarizing that. Yes, that's what I'm

saying.

MEMBER RITCHIE: And just to follow on, I mean I would agree with that. I think there are still some folks in the hospice community who are not doing so well on this measure, but by far and away, most are doing quite well. But we don't have the palliative care data and it seems like a great opportunity for us to have that data in the future.

CHAIR MORRISON: Last comments?

So I'd just go back to the preliminary, the work that NQF staff did, because we have to, this now comes up to voting. And the initial work by staff and the workgroup was that there was a moderate opportunity for improvement on this measure. That was the preliminary report which Rachel just put up, and we now come to the point where we get to use our clickers.

And what we are voting on is whether there is an opportunity for improvement on this measure and you get to vote under high which is 1, 2 which is moderate, 3 which is low and 4 is

insufficient.

We will take a vote. Please aim your clicks at Jean-Luc when he tells us to and not before, otherwise you'll be clicking all day.

MR. TILLY: Sure, you can go ahead and click now.

MS. ROILAND: And bear with us while there's a delay. We're just waiting for some remote participants to vote.

MS. JOHNSON: And while we're waiting, just a reminder, when you're thinking about gap, you can certainly look at the performance rates that are in front of you. You can think about whether those are high or low.

You can think about whether there's a lot of variation between the results. You can also think about whether there are any disparities in care, so any of those three ways would be ways to demonstrate opportunity for improvement.

Amy Sanders, we're still waiting on your vote. Have you sent something to Jean-Luc

or to Rachel? Amy, are you there? We're not hearing you.

OPERATOR: Amy has disconnected.

MS. JOHNSON: She may have stepped out. Let's just proceed without her vote.

MR. TILLY: Okay. And the results are 3 for high, 18 for moderate, 1 for low, and zero for insufficient. The measure passes.

CHAIR MORRISON: Well done, folks, your first vote. So we are going to move on to reliability and validity, and I'm going to go back. We're going to flip the order a little bit.

Ruth, can I ask you to talk about both? We'll start with reliability and then we'll move on to validity. We have to vote after each one, right? Let's just talk about reliability, she says.

MS. JOHNSON: And just a reminder too, under reliability this encompasses specifications and testing.

MEMBER MACINTOSH: Just a minute,

please. I apologize. I thought I was doing all of the screening and Cindi was doing the treatment, so --

CHAIR MORRISON: I can do it whichever way you would like.

MEMBER MACINTOSH: Well, why don't we let Cindi, why don't we do it that way?

CHAIR MORRISON: Cindi, why don't we start, we'll start with you. So we're on reliability.

MEMBER PURSLEY: Reliability specifications, they -- it was reliable initially, and now with the addition of, I think it's almost 4,000 hospices collected data, I think that it supports the reliability for hospice is there. I think the new data they presented supported reliability.

CHAIR MORRISON: Cindi, did you have anything? I'm sorry, Ruth, did you have anything to add?

MEMBER MACINTOSH: No, the reliability and reliability testing brings in to what the

developer said about changing the specification to no longer exclude hospice stays of less than seven days, and I have no problems with that. I agree with it.

CHAIR MORRISON: Open for discussion, rest of the committee. Christine.

MEMBER RITCHIE: Just as a question, are we, is a sample size of 20 acceptable for reliability testing for palliative care?

CHAIR MORRISON: And I turn to Rachel to answer that one.

MS. ROILAND: I'll take a first stab at it, and Karen, if you just want to correct me. We don't have specific standards for how large a sample size needs to be in any given testing situation, so it's your assessment of whether or not that sample size is sufficient.

MS. JOHNSON: And the only thing I would add is we would like to see testing that is representative, and that's in air quotes, to kind of encompass the kinds of providers that would be measured as well as the patients that would be

measured.

So you just kind of think about the scope, when you're thinking about rating just think about the scope of the sample. Twenty may be fine for some people, and twenty may not be anywhere close enough for other people so it's kind of up to you.

CHAIR MORRISON: Michelle.

MEMBER CAUGHEY: Thank you. That's the one I've been struggling with is that this is a measure that is in two different settings, one of which we have clearly reliable information and the other really hasn't been tested, and I'm not even sure that, you know, that the measurement that's coming hopefully will tell us something. This is going be true for a lot of our measures.

And so the assumption I'm making is based on the hospice data set data that it would be useful in the palliative care inpatient setting or even outpatient, ultimately.

And I just would ask whether that is what the other committee members are also

concluding?

CHAIR MORRISON: I've got Christine, then Paul, then I think Laura will probably respond.

MEMBER RITCHIE: Right, Laura. You can add this to your list of things to respond to. So this reliability testing with the 20 charts that were abstracted were, that was in one hospital, one facility?

CHAIR MORRISON: And Paul? Please.

MEMBER TATUM: I'll take my finger off. If one were to vote insufficient, I wonder if you could walk us through the implications to the rest of the process on that.

MS. JOHNSON: If the majority of the members voted insufficient -- sorry. If the majority of the members vote insufficient, then the measure would not pass reliability. We would stop discussion. The measure would go down. It would not be endorsed.

This one is, it is a little tricky because she has two different kinds of testing

but we're only asking for one vote. So you can use your own judgment about how you want to weigh and balance what you're seeing.

CHAIR MORRISON: Let me go to Woody, and then, Laura, I have the questions for you to respond to.

MEMBER MOSS: Yes, and so this is a clarification probably going to Laura. This is, I think, the 20 patients is inter-rater reliability not the overall sample, and at least in studies that I've frequently read, inter-rater reliability isn't just a sub-population it's the entire sample. So I'm just sort of commenting on that and seeing what Laura has to say.

CHAIR MORRISON: Dr. Hanson.

MS. HANSON: Oh boy, bunch of questions. So let me answer Woody's question first. So what we did with all the PEACE measures in the second data collection, second wave where we were purposefully looking at the hospital-based palliative care population, is that we took a random sample of the subjects who

were the target of those quality measures, a relatively diverse population, albeit within a single institution, and we basically started out with 20 records in order to see if inter-rater reliability was valid.

And when we hit a kappa, in this instance, of 0.89, we were satisfied that we had actually established inter-rater reliability. We did it in a rigorous manner in the sense that two independently trained raters were given the record to abstract and they abstracted all of the PEACE measures and then we compared their abstractions.

And so it was really the kappa, the level of agreement that on this quality measure led us to stop at 20. Could we have done more? We could have done more.

I think the second thing that I'd like to respond with, I think this is one of the fundamental challenges, honestly, for our field. And we face it in how to present the PEACE measures which were purposely, we were charged by

CMS to develop measures that would be relevant to the hospice population but also to the palliative care population.

The hospice denominator population is defined by who enrolls in the hospice benefit. It's easy to figure out. The palliative care population as we all understand is shifting sands. The denominator population for palliative care is changing every year.

We are trying to present these measures as they were originally specified which is to be measures that are relevant and applicable to the breadth of that population, while concurrently acknowledging NQF's dilemma which is that you'll need to see data that looks at populations in different settings with different characteristics in order to believe that a measure may perform similarly.

The gap may be different in different populations, but the measure itself scientifically may perform similarly across settings and across denominator populations.

I think for me as a measure developer, I would argue that these concepts of reliability are relatively intrinsic to the measure itself. There are some differences in the way the data is collected, so a little difference in implementation that may be important for reliability, but the specifications of the measure itself are the same for the two different data collection methods.

I don't envy you your job.

CHAIR MORRISON: So I have Paul.

MEMBER TATUM: Speaker phone. I can take my finger off and you can hear me if I lean in, I think. I'm happy to run through the algorithm. I find this very helpful.

CHAIR MORRISON: Paul, you've got to make sure -- there you go. Oh, still not close enough.

MEMBER TATUM: I'll move this forward so I don't have to spend the whole day fighting this thing. I'm very pleased to run through the algorithm. It's very helpful to me just noting

that the quantity under moderate can be low but the quality is moderate, and I think this metric is important.

And while we could waffle on the palliative side of things, that piece about the quantity of information being low to high is helpful for me to review that as we talk about the palliative numbers, and the hospice numbers are good.

I'm not supposed to speak to emotional advocacy for a metric, but I don't want us to keep throwing out baby with the bath water of measures that we desperately need.

CHAIR MORRISON: And just a clarifying question for both Karen and Laura. Laura, you had said that there would be data available on the palliative care performance coming forward, and that was my first question.

My second question to Karen is that, can the committee ask or make a strong recommendation that those data be presented back to us if the measure passes on, you know, on high

or moderate?

MS. JOHNSON: I think that would be really fair to ask Laura if you are at least considering presenting not just performance data but are you planning on doing additional testing?

For example, you've done signal-to-noise at the score level, are you planning on doing something similar to that?

MS. HANSON: Wasn't planning on it, but I'm here to listen. So I'm happy to take those kinds of inputs and recommendations.

CHAIR MORRISON: Anybody on the phone have a comment? I see Dr. Ritchie's card going up. Christine.

MEMBER RITCHIE: So this is just my, so please bear with me. This is my learning how to figure out how to support, effectively support sort of this process.

So I think nobody would argue that dyspnea is an incredibly important thing to measure, and the challenge, Laura, that you brought up is, of course, one that's very

significant and probably one we'll continue to struggle with throughout the day.

But there are two challenges that I see and I just wanted to get input from Karen and others about how to navigate these, so one is around the denominator.

So, if anything, just trying to figure out when we're talking about the denominator exactly what that denominator is would at least give me some sort of comfort about how likely it is that we're going to be able to replicate these kind of findings elsewhere.

And then the second is around the practice setting and this relates to sort of how data gets assessed through chart abstraction, et cetera, based on the setting of care. So, you know, if you have a certain kind of electronic medical record then your likelihood for being able to demonstrate reliability may be much greater than if you have a different kind of electronic medical record.

And so I'm struggling with those two

issues and how much to take them seriously or just to disregard them, I'm happy to disregard them but that's what I'm struggling with.

MS. JOHNSON: So let me start just quickly talking about the testing that they did for the hospital side, right. They provided one kappa value and we are assuming that that was for the numerator.

And it would be good, Laura, if you could just tell us for sure that you probably did some case finding or something and you couldn't really do a denominator kappa.

So I think the question for you guys is, do you feel like that without seeing data that you can make an assumption that one could accurately and reliably, actually I should say reliably, get that denominator knowing that different hospitals are going to pulling from different kinds of EHRs, it's just the world we're in, and so that's just something you have to balance there?

And I'm sorry, Christine, I forgot

your second question.

MEMBER RITCHIE: It's really about the challenge of engaging in reliability testing at one site when the reliability testing is predicated on how the charts are abstracted and what the electronic medical record, or not, looks like.

MS. JOHNSON: So here's the way I think about reliability testing. I think about it a little differently if I'm thinking data element versus score level.

For data element it's usually a small sample. And the idea with NQF is that it at least gives you a flavor of what can be done. It's no guarantee that every time you go out into the world that it will be similar at the data element level, but it at least gives you some comfort hopefully if the results are good that's reasonable.

The score level reliability, which they have not done for the palliative care hospital setting, the score level reliability is

actually a function of the variability and it's really more a function of the data that we're looking at.

So you can't necessarily say because the score-level testing in the hospice setting was high I know it's going to be in the hospital setting. You can't make that jump, but you can probably make a jump on the data element side. That's how I think of reliability testing.

So I'm not sure if that's answered your question or not. Are we getting there?

MEMBER RITCHIE: That's the last thing I'm going to say, or question I'm going to ask. So the data elements in the hospice setting is through the hospice item set, Laura, but the data elements are much more arbitrarily designated in the hospital because there's not a measurement tool that's consistent across all settings, right, or am I misunderstanding that?

MS. HANSON: For the hospital-based palliative care, well, really, hospital-based serious illness population some of whom had

palliative care, specialty palliative care, and some of them did not, we had a structured chart abstraction tool. We had training for the abstractors.

So I would argue that it was a process very similar to the HIS, very similar in its standardization for the abstraction of the data elements, the training of the abstractors, so that they could achieve inter-rater reliability because they had a standardized approach to the data collection itself. That's how the kappa of 0.89 happens.

It's not the same process. It's not the identical process to the HIS. When we collected that data, the HIS did not exist. I think it does raise a question moving forward, this convergence question around different forms of electronic medical records.

One of the things that we observed in the PEACE project which I found to be very interesting was that we asked for the collection of denominator data separately from numerator

data, so the data that would make you meet the numerator was separately abstracted in time on the same instrument but you couldn't tell as an abstractor that you were picking up the numerator and the denominator.

We tried out what it would be like to abstract, essentially, did this person with dyspnea get treatment within 24 hours? Not surprisingly, if you abstract the numerator and the denominator at the same moment in time you're much more likely to meet the quality measure.

So how the data is collected is really important. We tried in both the HIS and in the way that we collected the palliative care data to have a very rigorous and very standardized instrument in order to pull that data.

CHAIR MORRISON: I'm going to ask if there are any last, burning issues around reliability before we go to a vote. I recognize that this is confusing. I recognize that it is complex, and I recognize that there are no easy answers.

I do want to highlight Paul's comment that if you really get stuck, you have the algorithm which is extraordinarily helpful in terms of wading through this.

But I think we're ready to vote on the reliability criteria. And again 1 is high, 2 is moderate, 3 is low, 4 is insufficient. When Jean-Luc says go, point your clicker at Jean-Luc.

MR. TILLY: Yes, by all means, please go ahead.

MS. ROILAND: Bob, if you're on the line, if you could send your vote to Jean-Luc, I would really appreciate it. Thanks.

CHAIR MORRISON: Bob, if you're on the line and still with us, we need your vote.

MS. JOHNSON: And Bob, if you can hear us and you're trying to talk we cannot hear you, just so you know.

MR. TILLY: All right. The results are zero for high, 20 for moderate, 2 for low, zero for insufficient, so the measure passes.

CHAIR MORRISON: And then I think the

last thing just before we move on, as Karen reminded me that we would like to insert a comment that we would like to see some of the specs coming back, if that's okay, Laura, and even if it's not.

So we're going to move on to validity. And again just to very quickly highlight the aspects of validity testing, I'm going to do it just this time. You know, we're looking for the specifications. Is it consistent with the evidence testing, similar tests, how it was tested, and then the results and any threats to validity?

So let me ask, Cindi, you're doing treatments, right?

MEMBER PURSLEY: Yes.

CHAIR MORRISON: So let me ask Cindi to start with this one.

MEMBER PURSLEY: Validity testing, face validity and empirical, we have the 2015 data for hospice. It's extremely high, 1.2 million patients, almost 4,000 hospices. Again

for palliative care there's nothing available. I think that statistically hospice is extremely high.

Do you want to talk about threats?

CHAIR MORRISON: Yes, I was going to say why don't we talk about threats?

MEMBER PURSLEY: Okay.

CHAIR MORRISON: I'm sorry.

MEMBER PURSLEY: The developer has asked for an exclusion of the seven-day, patients with a length of stay under seven days be excluded. This is a very high population for hospices as we all know.

And in testing the validity it was found that it did not have an impact but it did offer a higher denominator for the data. So I think taking that out of this measure, that exclusion.

CHAIR MORRISON: Karen, do you want to just talk a little bit about face validity?

MS. JOHNSON: Sure. Just a reminder, for the palliative-care setting the developer did

do some face validity testing, and those results are there for you. And NQF does allow face validity as an indicator of quality as long as it's done at the score level which it was done.

And they also did some construct validity for that setting. We weren't exactly sure when we were doing the analysis from the staff exactly how that analysis was meant to demonstrate validity.

So one that might be something that Laura could help us understand a little bit more because with face validity moderate is really as high as you can go at least for that side, but the construct validity could be at score level and that might get you a different rating.

CHAIR MORRISON: Laura.

MS. HANSON: So for all of the PEACE measures for this sample of 500-plus patients in a single hospital site, my own institution, what we did for construct validity would typically or frequently be tested by an a priori hypothesis that says this measure measures this construct.

And I have another measure over here that also measures that same construct, so I'm going to collect both and then see how well aligned they are.

We didn't have anything like that to use for this quality measure, so as a consequence what we decided to do was a priori hypothesize that individuals with serious illness who screened positive for dyspnea would be more likely to meet the numerator condition for this quality measure, and that's true for the others in the set, if they were seen by a specialty team in palliative care than if they were not.

So we basically used that hypothesis as a kind of proxy for construct validity, saying this is a priority for the expertise of the interdisciplinary team and therefore we would expect the measure to be met more often.

In the case of this measure you can see that there was an incremental difference but not one that was statistically significant.

CHAIR MORRISON: Other comments from

the group? So just to remind everybody that the face validity only gets you to moderate under NQF rules, but if there's no more thoughts, I think we're ready to vote on validity.

Christine. Sorry.

MEMBER RITCHIE: So just a quick question about the removal of the exclusions. So my worry there is that if we remove the exclusions and we happen to have a hospice that's in a community that gets late referrals that this would be a potential disadvantage to them if they can't, if they're getting referral the day of, they will not be able to meet the measure.

And I just wondered, is there a plan to engage in a risk adjustment around that or how were you thinking about addressing that issue?

MS. HANSON: So the data's in the packet, but basically, very purposely, out of the HIS analysis for each of the measures where that seven-day exclusion had been utilized, the RTI analyst then provided a comparison to basically demonstrate that that does not preclude needing

the quality measure, that that is not as much of a threat to hospice practice as was assumed by the advisory panel when the measures were first specified with the seven-day exclusion.

And I have to say that there was a lot of back and forth about that initially in the PEACE measures because it would exclude so many hospice patients from being included in quality metrics. So we were relieved to see that the data didn't support that anxiety.

CHAIR MORRISON: Karl, last question.

MEMBER STEINBERG: Just wondering how difficult it would be to do 48 hours or 24 hours, you know, in relation to that concern.

MS. HANSON: So my RTI colleagues who are here can speak to that in even greater depth, but the reality is that we've been slicing the RTI data every way to Sunday and seeing basically that if you're going to meet this quality measure you're going to meet it through the initial assessment and management process.

CHAIR MORRISON: Everybody comfortable

to vote? Okay, back to Jean-Luc.

MR. TILLY: Great. So your options for voting on validity are high, moderate and low, and then insufficient number 4. Please go ahead.

Okay, the results are 1 high, 23 moderate, zero low, zero insufficient. The measure passes.

CHAIR MORRISON: Okay, guys, we have three more to go. Feasibility.

MEMBER PURSLEY: Okay, feasibility, preliminary rating was moderate. The data is incorporated into the hospice medical record and it's transmitted directly to CMS.

So I think that the feasibility for hospice again is high. It's a standardized item set and it's easy to pull the data. I think for palliative, again that's a question.

CHAIR MORRISON: Ruth, anything to add?

MEMBER MACINTOSH: The NQF group rated it moderate and I'm in agreement with that.

CHAIR MORRISON: Terrific, thank you. Open for discussion. Okay, I think we can move to a vote then. Jean-Luc.

MR. TILLY: Okay, great. Your options for voting on feasibility are 1 high, 2 moderate, 3 low, 4 insufficient.

CHAIR MORRISON: Rob, do you want to ask a question while we're waiting for the other three votes?

MEMBER SIDLOW: Sure. I'm sorry, this question might be a little late but it's a question about the actual specification. I was sort of digging in here and noticed that screening positive is anybody with a measure above zero on the dyspnea screening and that should trigger an intervention.

And just thinking practically, wouldn't a screen trigger be a little higher, like 4, because I wouldn't necessarily start an intervention at a 1, for example.

MS. HANSON: Welcome to my world. So a lot of debate back and forth about that

specific aspect of specification, I'm amazed no one else raised it so I'm glad you did.

And it is still an issue that's under active discussion in thinking about implementation of this quality measure and the one that we'll eventually discuss on pain assessment which relies on that same concept that you screen positive but at what level of severity?

The difficulty is that our advisory panel could not come to consensus that there was a clear-cut point that should be used for dyspnea, so we ended up with a concept that if the patient said they were short of breath, sort of going back to the basic concept of what is a symptom, if they said they were short of breath they were short of breath and it needed to be addressed.

It's important to recognize that treatment did not require administration of medication. Treatment could be any medication or non-medication based approach that addressed the

symptom of dyspnea.

But I do agree with you that it's a point that I have questions about as well. We used the specification that our advisors led us to use.

CHAIR MORRISON: Jean-Luc, can we have a final? And then I think Cindi has a comment, I know.

MR. TILLY: Actually we're missing one vote in the room, if you all could just try it just once more with the clicker. That's right, yes. Press as many times as you like.

MS. ROILAND: And remember to point it at Jean-Luc. Okay, we've got 24.

CHAIR MORRISON: Why don't we do that and then I'll take Cindi's question.

MR. TILLY: Very good. The results are 1 for high, 23 moderate, zero low, zero insufficient, and the measure passes.

CHAIR MORRISON: Cindi, you had a comment?

MEMBER PURSLEY: Yes, I just wanted to

say as a reminder that it's also the patient's decision as to whether or not their dyspnea needs treatment, because we may have patients that appear very dyspneic to us but they're very comfortable with it and prefer not to treat. So I just wanted to throw that out there.

CHAIR MORRISON: Yes. Let's move on to usability, last voting -- no, second to last voting.

MEMBER PURSLEY: Usability and use, these measures are not publicly reported. My understanding, it's going to be middle of 2017 at the earliest that these will be publicly reported.

Accountability use or planned use, it's very much a part of what hospitals are tracking as well as CMS, so I think that the usability and use, let's see, was moderate and that was the preliminary rating.

CHAIR MORRISON: Comments? Actually, I'm sorry. Ruth, anything to add? You were the other discussant.

CHAIR MORRISON: Nothing to add.

CHAIR MORRISON: Comments, thoughts? Open to vote. Jean-Luc.

MR. TILLY: All right, voting is now open. Your options are 1 for high, 2 for moderate, 3 for low, and 4 for insufficient.

So we're looking for just three more votes in the room, so if you have some doubt about whether or not it went through just try pressing it again. That's exactly right. A light or a flashing number, really, is a signal that it succeeded.

CHAIR MORRISON: Make sure everybody points at Jean-Luc.

MR. TILLY: I think we need just one more, sorry. Okay, great. So the votes are 2 for high, 22 for moderate, zero for low, and zero for insufficient information. The measure passes.

CHAIR MORRISON: So guys, the long primary season is over. We now move into the general election. This is where you get to vote

on up or down on endorsement or not.

MR. TILLY: All right, so for overall suitability for endorsement, there are just two options, one yes and two no.

MS. ROILAND: All right, we needed to meet that magic 24 number and we need one more in the room, so if you all could point again. Oh, are we -- oh, he's out of the room. Okay. All right.

MR. TILLY: Thanks very much for mentioning that. The results are 21 yes and 2 no, so the measure passes.

CHAIR MORRISON: All right, guys, just before we break, we need a quick look at whether there are any competing measures and there's opportunity for the committee to talk either about competing measures or harmonization.

There are two potentially competing measures identified, but not really. I don't, you know, they're up there. One, you could see them if anybody has a comment about that.

Okay, so we're all done, folks. We

have 16 more hours if we continue this pace just on measures alone. Fortunately, Deborah's going to read us through things after the break, and she's a social worker and handles family meetings much better than I do.

So let us reconvene at five of 11. We will try and get ourselves back on track. And just a heads up, after this we're going to be moving much more quickly. This was deliberately slow to get everybody onboard.

(Whereupon, the above-entitled matter went off the record at 10:38 p.m. and resumed at 10:55 a.m.)

MS. JOHNSON: Okay. We need to go ahead and get started again. So if you'll take your seats please, we'd appreciate it. We are running a little bit late now, but I think we can catch up with our timing. So, Deborah, I'll hand it to you.

CHAIR WALDROP: Okay. Welcome back, everybody. And first, I want to welcome Amy Berman. Thanks for joining us. And we need to

ask you to just briefly introduce yourself, and then state if you have any disclosures on the measures.

MEMBER BERMAN: Good morning, everyone. My name is Amy Berman. I'm a Senior Program Officer at the John A. Hartford Foundation in New York City, which is a roughly $500 million dollar foundation focused on the improving the care of older adults.

A good portion of the work that we do is in the field of palliative care. I've been Diane Meyers' program officer, for example, for the past 11 years. We also tend to find many other things that go on in this field. So this is also, professionally, my background around improving quality of care for people who are seriously ill and at the end of life.

I also, five and a half years ago, was diagnosed with Stage 4 inflammatory breast cancer. The prognosis, 11 to 20 percent survive to five years. I'm a person who made choices and publicly speaks about them and writes about them.

And because of those choices, I've lived very well, very well. And I feel very well, thanks to the work that many of you in the room do. So I thank you all personally and professionally for what you do.

In the interest of disclosure, should I continue into the disclosure piece? The only disclosure that I have to report is that I was on the ASCO Palliative Measures Panel. And that was this past year.

And the measures that we looked at were 209, 210, 211, 213, 215, and 216. So on those measures I will not be weighing in. I will be recusing myself. And I will not be providing comment. Thank you.

CHAIR WALDROP: Okay. Thank you, Amy. So now we will move into a little quicker pace for the next several measures, for the next, rest of the measures for today. But I have to ask your forbearance. This is new for me. And so, I'm going to just jump in, and hope that I can keep the process moving, with the assistance of

Karen, Rachel, and Sean.

So, we're going to move on now to Measure Number 1639, which is a paired measure with the measure we've just previously reviewed. This is Dyspnea screening. And I'm going to turn to our discussants, and ask, I believe it's Ruth who will be starting with the evidence for us, please.

MEMBER MACINTOSH: Yes, thank you. This should move quicker since we spent the time last time. The evidence is, first of all, this is number 1639, Hospice and Palliative Care Dyspnea Screening.

The percentage of hospice or palliative care patients who are screened for dyspnea during the hospice admission evaluation or the palliative care initial encounter, it is the level analysis clinician, group practice, and facility.

It's not a new measure. It's a maintenance measure process. And the evidence is there from previous. And the developer presented

updates in evidence. So the evidence is there. I don't know if you need a vote.

CHAIR WALDROP: No, that's okay. Cindi, do you have anything that you'd like to add?

MEMBER PURSLEY: I agree.

CHAIR WALDROP: Okay. Discussion, questions, comments from the committee? No comments from the committee. Any comments from the committee?

MS. JOHNSON: So this is Karen. I'll step in here as staff. You'll note that in our preliminary analysis we had selected insufficient. And that is because that when we were reading this we felt that there was not strong evidence to show that screening actually improves outcomes.

So, I'd like to invite you to discuss this just a little bit. Is there additional data that might show this? Or, if not, I can talk about a pathway that we might could go down if you do feel that the evidence is actually

insufficient.

Again, we are looking for, in terms of evidence, we're looking for a link to show that screening improves outcomes. That's the link that we're looking for.

CHAIR WALDROP: Okay. So, comments on whether or not the evidence is sufficient on whether screening makes a difference in treatment? Thanks. Michelle.

MEMBER CAUGHEY: So, clearly there's not the evidence presented, nor am I aware of it in the literature. But I think this is a measure, like the pain measure, that should be considered by exception.

CHAIR WALDROP: Other thoughts? Comments?

MEMBER MACINTOSH: The developer did add two guidelines. And one was 2011 British Columbia Medical Services. That was based on a systematic review of the evidence. But the evidence wasn't graded.

I agree that this should be by

exception. We can't treat dyspnea without the screening. And we have a roads to go with the palliative care, as discussed before.

CHAIR WALDROP: Doug.

MEMBER NEE: I just concur with that last statement. Absolutely.

CHAIR WALDROP: Thank you.

MS. JOHNSON: So, let me explain how we would do this. If you agree that there isn't evidence showing that link, but you still feel that there might be room in the NQF portfolio to continue to endorse this, we now have what we call an exception to the evidence criterion.

And we talked about it here in the preliminary analysis, that if you look at your algorithm, and you guys have the algorithm there in front of you. I'd invite you to pull that out. And let's just walk through the algorithm very quickly, thank you, Marcia.

So, the first box, Box 1, asks if it's an outcome measure. And we know that it's not. It's a process measure. And in Box 3 we want to

know if it, the measure is based on a systematic review and grading of the body of evidence.

So, while there were systematic reviews included here, they were not graded. And we did not see a summary of what we call the quantity, quality, and consistency in the body of evidence. I think the boxes that were checked up there may have been checked erroneously.

That takes us to Box 7, is empirical evidence submitted, but without systematic review and grading of evidence? That's the idea that if there are articles, but there's been no systematic review, that's okay too, as long as you're pulling the whole body of evidence, okay.

But I don't believe that additional articles were submitted. So that takes us down to Box 10. And that's the path for evidence exception. So, basically, what this asks is, are there, or could there be performance measures of a related health outcome, or evidence based intermediate outcome or process?

And if the answer is yes, then we

would not want you to rate this as insufficient with the exception. However, if they're not, then you would go on over to Box 11 and ask if there's a system, if there's evidence of a systematic assessment of expert opinion.

Now, in this case -- and sorry, that the benefits of what is being measured outweigh the potential harms. In this case the answer would be yes, because you have the systematic reviews that just weren't graded, right.

And then finally, that would take you to Box 12. Do you, as the steering committee, agree that it is okay to hold providers accountable for performance in the absence of empirical evidence of benefits to patients?

And you would follow your thinking there. If you think it is okay then you would rate as insufficient evidence, but with an exception. Otherwise, you would rate as insufficient.

So, when we get to this voting, and I apologize, my glasses are not as good as I would

like. You have four options here. So, if you believe that you would like to go down the exception path, what we need for you to do is to vote insufficient on this voting activity right now.

Thank you, Jean-Luc, that helps. If more than 60 percent of you vote insufficient, then we will have a second vote as to whether we will actually invoke the evidence exception. And that'll be a yes or no, okay.

If you believe that what is here is actually sufficient information, then you would vote high, moderate, or low, okay. So you see the difference? It's a little tricky, so I want to make sure everybody's clear. Arif.

MEMBER KAMAL: I'm just going to submit here at the eleventh hour that I knew there was a paper. I'm just working to find it.

MS. JOHNSON: Okay.

MEMBER KAMAL: From a book chapter we wrote actually. The first author is Seow. It's from the Journal of Oncology Practice from 2012.

And what they show is that as documentation of both pain and dyspnea went up, clinical decisions to address that symptom went up in a dose response way.

MS. JOHNSON: Okay.

MEMBER KAMAL: And so they pointed out both for pain and shortness of breath. So as documentation went up from 37 percent to 71 percent, dyspnea-related actions from the clinician went up from 4.2 percent to 37 percent. So, screening led to clinical action.

MS. JOHNSON: Okay. So that is a case where we actually have some additional evidence. And that's okay. Because you guys are the experts. And you would know this. So, you can certainly take that into account, what Arif has told us about. Any other questions about process? Go ahead.

MEMBER HANDZO: Yes. I would just ask then, help me out here on the algorithm. Run -- assuming that we knew about that, and it was included in the evidence analysis, and the staff

knew that, where would that lead you in terms of a recommendation?

MS. JOHNSON: If it were there, and if it were summarized in a way that I could understand it, right, and if I also felt that that was pretty much the body of evidence. So that there's not other things out there that would show something different.

And I felt like the whole body of evidence is there, then I would probably say that I would probably land on, if I were doing it, I would probably land on moderate as opposed to low. If I thought that it wasn't the full body of evidence, then not so much. If I thought that there was conflicting evidence, I would land on low.

MEMBER NEE: So we're back to square one, based on what you just said.

CHAIR WALDROP: Okay. Okay. So, I didn't see the order of these. But I'll go with Gregg first.

MEMBER VANDEKIEFT: So, I'm curious

where that leaves us, since this is just brought up. You know, we haven't had a chance to review that evidence. And even, Arif, I'm not sure that documentation equals screening. So there's an inference there. How do we integrate that new information into our decision making today?

MS. JOHNSON: I'll take it, and kind of do the, my favorite answer at NQF. Is it depends. It really depends on how comfortable you are with making that choice. If you feel like that that isn't the link that you want to see, then you would vote accordingly.

If you feel like that is enough of a link to make you feel like that the screening would lead to better outcomes, then you could vote the other way.

I will say that if it turns out that we go a route other than exception, I would ask Laura to go back, and just add that information to the submission. So that it's very transparent that if you did land on something different, where that would come from.

CHAIR WALDROP: I think what I'm going to do is take the committee members first. And then I'll ask Laura -- offer Laura a chance to respond. So, Paul?

MEMBER TATUM: Mic has moved forward so you can hear me when I hit the button. To confirm, insufficient kills the process, but clarification about a low evidence, the process moves forward?

MS. JOHNSON: So, clarification. If a high majority, that is more than 60 percent of you, land on insufficient, it doesn't kill it yet. It would take it to the next question, which is, do you want to actually invoke the evidence exception? And we would have a vote on that as well.

MEMBER TATUM: My apologies. You were clear on that earlier, and I used the wrong terminology. But the low continues the process, correct?

MS. JOHNSON: If a majority land on low, the measure dies.

MEMBER TATUM: Got it.

MS. JOHNSON: So there is a very large difference between low, which would probably suggest conflicting data, something like that, as opposed to insufficient, which is, there's not enough to actually let me rank it one way or the other.

CHAIR WALDROP: Steve, and then Deborah, and Michelle. Deborah.

MEMBER WIEGAND: I just had one question about the article that you just reported. I'm sorry, I don't know his name.

MEMBER KAMAL: Arif.

MEMBER WIEGAND: Arif. Just a question about the article. What population -- thank you.

CHAIR WALDROP: Michelle.

MEMBER CAUGHEY: While it's great that we have introduced the new article, and some new information, we don't know that that respects the process that we've all undertaken, which is to review what we had in great detail, particularly

those that were responsible for the measure.

And so, I would suggest that we take that under review for our next panel. And that I don't think it should be added into the discussion, my apologies to my colleague.

CHAIR WALDROP: Paul, your card is still up. Did you have another question?

MEMBER TATUM: Just pleased to announce that this is a great process. Thank you.

CHAIR WALDROP: I need to apologize to our developer, Laura Hanson, for not including you to begin with, but I'd like to give you the chance to address some of these questions.

MS. HANSON: No apology needed. I'm delighted to be skipped over. It's wonderful. And I feel like I introduced the pain screening measure because it's a paired measure with the, I mean, I'm sorry, the dyspnea screening measure, because it's a paired measure.

I mostly just wanted to comment back to your question, Gregg, that if the committee

decides, and there may be a procedural point here, to include the Seow reference, these process measures are based on documentation. So that's how they're, that's the body of evidence that screening actually occurred. So in that sense it would parallel the reference.

CHAIR WALDROP: With due respect for the -- oh, Deborah.

MEMBER WIEGAND: So just to ask a process question. I thought though if there was evidence that hadn't been presented, that we were supposed to bring it up in today's discussion. Is that not correct?

MS. JOHNSON: It becomes a little tricky. Because we want you in one way to look at what's in front of you, and vote that way. But we also know that you bring your own set of expertise. And you might know something in and of yourself, or your colleagues may tell you something that may change a little bit.

So all we can really advise is for you to take what you've heard, and do the best that

you can with what, how you understand our criteria.

CHAIR WALDROP: Okay. With due respect for the process, I'd like to move us along, unless there are any other comments to vote, to vote on the evidence. Jean-Luc, let us know when you're ready.

MR. TILLY: The polling is now open for a vote on evidence. Your options are one high, two moderate, three low, and four insufficient.

(Pause.)

MS. ROILAND: Amy, if you're on the phone, if you could text your vote to Jean-Luc, we'd appreciate it. Thank you.

MEMBER SANDERS: I'm trying.

MS. ROILAND: Or if the chatbox is better, you can use that too, or email.

MR. TILLY: The votes are zero for high, three for moderate, one for low, and 19 for insufficient. So the verdict on evidence is insufficient.

CHAIR WALDROP: So then we move to a vote about the exception.

MS. JOHNSON: And let me add just here with this vote we actually need for this to carry greater than 60 percent. So, just so you know, the implications of your vote are.

If we, if more than 60 percent say insufficient evidence with exception, it will pass this criteria and we will continue. If not, there is no gray zone for this one. So if not, then it ends here.

MR. TILLY: So the vote for evidence on the potential exception to empirical evidence, your options, there are two options, one insufficient evidence with exception, and two no exception.

(Pause.)

MR. TILLY: The results are 22 for insufficient evidence with exception, and one for no exception. So the measure passes with insufficient evidence with exception.

CHAIR WALDROP: Okay. Thank you. So

we'll move into looking at gaps in care, and opportunities for improvement. These will address disparities in this section. And I'm going to go back to Ruth and ask if you would like to lead us through that?

MEMBER MACINTOSH: Yes. For the facility level for hospice we have quite a good data source from the same amount of the HQRP for 2015. And we have high percentiles for the hospice. We don't have specific data, same thing for palliative care yet.

It was rated low opportunity for improvement, because most hospices are reporting on this measure. The performance seems to be tapped out. However, in disparities there is not as quite the high percentages there.

And again, this gets to low opportunity for improvement, but opportunity for both the disparities, and also the actual dyspnea treatment percentages we saw.

CHAIR WALDROP: Any questions or comments from the committee? Okay. I believe we

are ready --

(Off microphone comment.)

CHAIR WALDROP: Did you have a comment on the disparities, Christine? No? Okay. Seeing no further discussion, I believe we move to a vote on gaps in care. Jean-Luc?

MR. TILLY: Yes. The polling is open for the performance gap vote. Your options are one high, two moderate, three low, four insufficient.

(Pause.)

MR. TILLY: So the results are zero for high, 11 for moderate, 12 for low, zero for insufficient. So the measure is in the gray zone where consensus is not reached.

MS. JOHNSON: And just as a reminder, what that means is it did not pass this criterion. However, since it's in the gray zone, we will continue to discuss the measure.

CHAIR WALDROP: Okay. So we'll move on to reliability. And I just need to clarify, do we, we will vote on reliability and validity?

MS. JOHNSON: Yes, we will. This measure is quite similar to the one that we've just discussed. So the methods will be exactly the same I think. And all of your measures, Laura, right?

What will be different is the results. So, because the results are different, I would ask that you vote on this one after any discussion you may want to have.

And remember, for reliability it's not just the testing. So you're not just thinking about the testing results. You're also thinking in this portion of the discussion about the specifications. So if there's any questions or concerns you have about specifications, now would be the time to bring those forward.

CHAIR WALDROP: We start with Ruth, and ask if you have anything you want to add about the reliability and specifications.

MEMBER MACINTOSH: No. As Karen said, it was the same as last, as far as the actual testing. But for, with the results in the

facility level -- let me just look here.

The clinician, group practice level for palliative care, the single kappa value was 0.91. So the raters agree 91 percent of the time over. And for facility level the split-half analysis was .83.

The developers reported a signal to noise ratio of 0.98, with one being the highest. The group for NQF rated the guidance from using the reliability algorithm for the facility level as high, and the clinician level using that algorithm as moderate.

CHAIR WALDROP: Thank you. Any comments from the committee? Laura? You had your thing up.

MS. HANSON: Oh, I apologize.

CHAIR WALDROP: Oh, okay. No worries. No worries. Okay.

(Off microphone comment.)

CHAIR WALDROP: Seeing no other indication of discussion, I think we're ready to vote on reliability of this measure.

MR. TILLY: The polls are open for voting on reliability. Yes, this is a new vote. Please hit one for high, two for moderate, three for low, and four for insufficient.

All right. We're missing actually just one vote in the room. So I'm sorry to make you keep doing this. But if you could try just one last time. Pointing it over here. Thank you.

All right. So the results for reliability are, four high, 19 moderate, zero low, zero insufficient. The measure passes reliability.

CHAIR WALDROP: Thank you. We'll move on to validity testing. And I'll go back to Ruth and ask if you have anything you'd like to add.

MEMBER MACINTOSH: I think the group, the developer noted that the specification was consistent with the evidence. The face validity is the same as what we discussed before for the palliative care. Same with the construct validity.

For hospice, they used the full year '15 data for almost 4,000 hospice organizations. And the validity testing results were, correlation results were positive and statistically significant for the facility level.

But they did find the magnitude of correlation was lower than expected. And that, the developer states the reasons for this may be due to clustering of scores, skew distributions, and low score variability.

Split out for the palliative care, the face validity indicated broad endorsement of the face validity of the measure from the stakeholder group. And the, it was felt that the exclusion was in there.

And we discussed that last time for removing the previous length of stay. We've got the assurance that when it was less than seven days it didn't greatly impact. So the exclusion's consistent with the evidence.

And the meaningful differences, the development -- developers examined 95 percent

confidence intervals to determine the proportion of hospices with results significantly different from the hospice level mean. And it indicated that 36.7 percent of hospices had a quality measure score that was significantly different from the national mean.

There was a low rate of missing-ness. And so, using the guidance from the validity algorithm, this was rated as moderate for both.

CHAIR WALDROP: Okay. Thank you. Any comments, or questions, or thoughts from the committee? I think we'll move to vote on validity of this measure.

MR. TILLY: All right ,the polling is now open to vote on validity. Your options are one high, two moderate, three low, four insufficient. All right. The results are two for high, 21 for moderate, zero for low, and zero for insufficient. The measure passes on validity.

CHAIR WALDROP: Okay, thank you. We'll move on to considering feasibility. And I

realize I've been leaving Cindi out. So I should turn to Ruth and Cindi, and ask if either of you have thoughts about -- like -- share about the feasibility of the measure.

MEMBER MACINTOSH: No. We consider it feasible. The group rated it as moderate, because we have the electronic health records for the hospice. And we realize that there can be a substantial collection effort for the palliative domain.

CHAIR WALDROP: Okay. Thank you. Any comments, thoughts? My understanding from Karen, if I can articulate this correctly, is that because of the similarity of our vote last time, we can decide to just carry over the vote from our, the last measure.

Can I see the thoughts of the committee? I see a thumbs up. I see a couple of thumbs up. Maybe you can explain it better anyway.

MS. JOHNSON: Sure. The question for you is, is the feasibility for this measure any

different in your mind than the feasibility of the last measure that we had?

If not, if there's nothing new, we will allow you to carry over votes from the last time. If anybody objects to that, we will have an actual vote, but --

CHAIR WALDROP: So, let me just ask if there's anyone who objects to carrying over the vote? Okay. Then I don't know what the process is. But we just apply the same number of votes to this measure. Okay. Which moves us on to usability and use of Measure 1639. And --

MEMBER MACINTOSH: This is not, it's the same as last. It's not currently used for public reporting for the hospice. It is used in the accountability program. It's the exact same as last time, if we're allowed to carry over the vote from that.

CHAIR WALDROP: Let me first ask if there are any comments about usability and use on this measure? And then I will invoke the same statement, which is that we can carry over our

vote, unless there's any significant difference that anyone notes between this and the previous measure.

Does anyone object to carrying over the vote? That said, brings us to the end of Measure 1639. We will look at competing measures?

MS. JOHNSON: No. We do need to do an overall vote for suitability.

CHAIR WALDROP: I'm sorry. We do need to do an overall vote for suitability of this measure. So, Jean-Luc, when you're ready, let us know.

MR. TILLY: Yes. So we're voting for overall suitability --

CHAIR WALDROP: So, we have a question before we vote. Sorry.

MEMBER RITCHIE: Just, again, as a point of clarification. So we didn't receive consensus on one of the measures, I mean, one of the criteria. How does that influence this overall voting?

CHAIR WALDROP: Good question. I'm going to ask Karen if you can clarify that for us?

MS. JOHNSON: So, this is where you really get to weigh individually how much you weight that one vote in yours specifically. So, if you voted low on reliability, that might be enough for you to not want to vote suitable for endorsement at all. And you would vote that way.

But you may say that in balance of what we know about evidence, what we know about gap, et cetera, you would still let that go through. And you could vote yes for overall suitability. So, it really is how you balance all the criteria going forward.

And I'll stop there and see if Elisa or Marcia had any other -- maybe you can put it in a little bit different way? Yes, clarification, Elisa.

We had consensus not reached on, was it reliability? Do we remember? I think it was. If they vote, and they vote something other than

consensus not reached --

CHAIR WALDROP: Gray zone was gap, sorry.

MS. JOHNSON: Gap, okay.

CHAIR WALDROP: It's gap. Yes.

MS. JOHNSON: Sorry. My mistake. It was gap.

MEMBER RITCHIE: And the reason I think that's important is because it raises an important question that will be a question that we keep asking, as a question some of us were discussing over the break, which is, how high is topping off?

And is there a topping off for a measure that's considered to be a basic, fundamental aspect of service and care? So that's just a question that I had to bring to the group.

MS. MUNTHALI: So, because you didn't reach consensus on a major sub-criterion, we will hold the overall suitability vote until the post comment call.

MS. WILSON: And if I could speak to topped out measures, This is an issue that comes up in every single committee. So this is a familiar conversation. Often, how high is high? The committee has to decide that.

But I think one of the larger issues is, if we declare a measure topped out, and it were eligible for reserve status, which you heard Karen explain earlier, does that mean we stop paying attention to it?

And that's often at the crux of the debate, is if a measure is topped out, and it goes to a reserve status, will people stop paying attention? And often the conversation falls into two camps, which is, we don't want it to go to reserve status, because we're afraid people will ignore it.

And at the same time there's that balance that's if it's topped out, should we be, should not us, NQF, the royal we, should we be requiring providers still to report it? So, it's a challenging issue, but it comes up in every

committee.

CHAIR WALDROP: And we're not voting at this point? We'll wait until the post --

MS. MUNTHALI: Yes. You're not going to vote. What we're going to do is wait for the public and member comments to come in, and see if that helps to move you one way or the other on the areas in which you weren't able to reach consensus. And then you take a final vote then.

CHAIR WALDROP: Thank you.

MS. JOHNSON: So, apologies for leading you wrong on that one

CHAIR WALDROP: So, does that conclude this measure?

MS. JOHNSON: So that concludes this measure. So, Laura, you're off the hot seat for now.

CHAIR WALDROP: Thanks, everybody. So that moves us on to consider 0209. but I just want to remind us that we have four measures and a discussion about related and competing pain measures to consider before lunch. So, I'm going

to try to move us a little faster. Okay. So, 0209.

Okay. So, Doug and Debra are our discussants for 2000-0209. But I think I'd first like to start with asking our developers if you could just introduce this measure for us briefly.

MS. SPENCE: Thank you. So, let me start with introducing ourselves first. I'm Carol Spence. I'm Vice President for Research and Quality at the National Hospice and Palliative Care Organization. This is my colleague. Carl, do you want to introduce yourself?

MR. SCHEFFEY: My name is Carl Scheffey.

CHAIR WALDROP: Make sure your mic's on.

MR. SCHEFFEY: Ah, okay. My name is Carl Scheffey. I'm the Director of Analytics at NHPCO.

MS. SPENCE: Okay. So very, very brief introduction to this measure. As Laura

did, I'm going to just tell you a little bit about its origins. It goes all the way back to 1998 when NHPCO convened, along with other organizations, something we call the Outcomes Forum.

Based on then, back then if you all remember Pathways? Well, we had developed a pathway for end of life care. And based on that pathway, were created by this Outcomes Forum three what we called outcome measures. But they were at the, you know, sort of up here at the 20,000 foot level.

And so part of that then became, look, there's three outcome measures for effective breathing, self-determined life closure, and safe and comfortable dying. So there's where that comfortable dying piece comes in with this measure.

There's been a little confusion this measure was meant to encompass all of comfortable dying. And that is not true. Because this particular measure was what then was called an

instrumental measure.

In other words, one piece of that comfortable dying picture, which the comfortable dying was actually termed an end result outcome measure. And this was one instrumental measure in support of that. So that has been a little bit of a point of confusion, you know, along the way.

So, this, again, this measure was developed as an outcome measure, patient report outcome measure, as you know. The data are collected in two phases.

Initially on admission to hospice patients are asked a yes/no question, are you uncomfortable because of pain? And for those patients who answer yes, that they are uncomfortable, they then go into the measure.

And then on follow-up within, between 48 and 72 hours there is a check back with those patients to ask if their pain was then brought to a comfortable level. So those two pieces of data collection are what constitute then the, this is

where you get the numerator and the denominator, you know, for the measure.

I want to say a little bit about the benefits of this measure. It is the patient's voice. You know, you go on to, once those questions are asked, then the clinician, the hospice clinician goes on to do their full assessment with whatever assessment tool is appropriate for that patient.

So, you know, as you probably know, pain is subjective. So, for example, on a zero to ten scale, one person's two could be another person's seven. This measure allows for the patient to simply say, indicate are they comfortable, or are they not comfortable.

It doesn't presuppose or impose anybody else's, including the clinician's judgment on what level of pain, when you're looking at intensity, or anything else, you know, needs to be achieved, you know, for that patient. And I think that is one of the primary benefits, you know, of this measure.

CHAIR WALDROP: Thank you.

MS. SPENCE: So that, I'm just, yes, I think that's pretty much. I can also, again, a little bit about, a little more about history or not. But just very quickly, I think again one possible point of confusion is, so this measure is, has had NQF approval, and was one of the measures that CMS chose to implement for hospices during that very first year of quality reporting that was required, you know, for hospices.

CMS then chose, after not even a full year of data collection, to drop the measure, and move on to the process measures that are now in HIS.

I would think about, I would also want to give you a little bit of a context for that, so that you understand some of what was going on in the hospice community. And perhaps, part of the decision making on CMS's part reflected this.

Hospice was one of the three provider types that was pinpointed in the ACA legislation to institute a hospice quality reporting program.

Because we did not have one before that. Hospice is coming very, very late to the table in terms of looking at quality.

Again, this measure was put in place back in 1998. But there wasn't even a conditions of participation requirement for a quality assessment performance improvement program for hospices to have their own until 2006.

So, hospice's thinking, the individual provider's thinking about quality, what constitutes a quality measure, what's a numerator, what's a denominator, I fully admit that my hospice community was rather far behind, you know, the times in that.

So this was the first pre-specified measure that hospice has had to implement, unless they were voluntarily using this measure. And being a patient self report it meant data collection in a way that they were not accustomed to. And it was tough going, you know, to be absolutely honest about it.

And also, it was the first time that

they had to put together a system for collecting data that they then had to turn around and submit to, you know, to CMS. So that, in and of itself, was a huge step for hospices.

So I would say this measure, you know, had challenges to start with. I would've really liked to have seen it given, left in place a little bit longer. Although I do understand CMS's decision.

Also though, it was my understanding from CMS that they at some point were interested in revisiting this measure, and working with us as the developers to perhaps, you know, see whether we could, you know, modify it or do some other things to make it, you know, get it back out there with hospices. Okay.

CHAIR WALDROP: Thank you for the context. So I'm going to turn to Doug and Debra. I don't know if you would like to take the primary discussant role. Okay.

MEMBER NEE: Sure. I'll do that. Okay, so as far as evidence goes, essentially

pain is subjective. The evidence stands. There was new evidence that was provided. It really was an article out of the Oncology Nursing Forum in 2002.

Just, investigating the symptoms of distress and quality of life in with patients with cancer, and newly admitted to hospice care, did find a strong relationship between pain and distress. At this point in time, I don't really see much need for re-discussion or repeat voting.

CHAIR WALDROP: Thank you. Comments from the committee? Consideration of the context and of the statement that we don't maybe need to reconsider evidence. Paul. And then --

MEMBER TATUM: I would like to speak to the opportunity to modify this going forward.

CHAIR WALDROP: Okay.

MEMBER TATUM: And I'll provide anecdote from one hospice regional number, as opposed to national numbers. I think it's a great measurement. I think we need this measurement. My concern is that I think it ought

to be everyone in the denominator. And I'll give an anecdote.

In one month's QAPI meeting, only 9.8 percent answered they had pain and got included in the measure. Yet 96 percent of the admitted patients had an opioid.

Because, in the Midwest, maybe somebody coming to your home you're being socially polite and saying, no, I'm fine. Thanks. Or maybe they knew a nurse was coming, and they planned ahead and took the pain medicine before they arrived.

And so, when the question was asked, at this moment are you uncomfortable? You're uncomfortable. So, just as you move forward, I'd like to encourage making that measurement all hospice patients on admission.

CHAIR WALDROP: Thank you. My understanding is we're not able to redo specifications of a measure at this point.

MEMBER TATUM: Oh. I just heard her comment that was an invitation from the future.

So I apologize if that was inappropriate.

CHAIR WALDROP: No. Do you want to comment on that? Okay.

CHAIR MORRISON: Just a comment that, if we're going to get through the day, we're not going to be able to really focus, drill down on the specs of the measure and redoing the measure that's presented.

MS. SPENCE: No. If I could just respond? Because I think I can clarify a little bit. That again, this measure was meant to focus on those patients, not who had their pain managed or, you know, or had pain. But who were uncomfortable on admission.

Therefore, it was an impetus to get busy for the hospice, and address that. It, you know, so it is very possible -- so the person who says, no, I don't have pain on that initial assessment, doesn't mean that you don't, that you ignore the fact that they're on opioids, et cetera.

It's just you don't, it's the

addressing the person who is in pain, was intentionally the focus of the measure. With the understanding not everybody who was on pain medication was going to be uncomfortable on admission. So --. Absolutely.

CHAIR WALDROP: Thank you for the clarification. Cindi.

MEMBER PURSLEY: I am also addressing this from a user standpoint. When you ask patients if they have pain they frequently have arthritis. They have neuropathic pain from their diabetes. They have pain from their cancer.

And you may be able to manage that cancer pain, which is why you have them on hospice. And yet, when you ask them at 48 hours if their pain is under control they may say that, well, no, my joints still hurt.

So it's, you're not generally addressing just a pain in a place. And so, it does make it a little bit difficult sometimes to utilize this question in a way that's meaningful.

CHAIR WALDROP: Okay. Thank you.

MEMBER KAMAL: I'm just going to really quickly say, so, you know, in palliative care we don't have a lot of outcome measures. So it's important that, you know, as a field we sort of keep an eye towards maintaining the few that we have.

From a policy perspective we, you know, we really can't participate in reimbursement mechanisms, in Part A or Part B, without having outcome measure as part of our overall cohorts. I think we should keep an eye towards that.

Because there's not many that we're going to review today. So, I know from an NQF assessment perspective, outcome measures are, you know, reviewed at a higher level, which they should be.

I think from a clinical content expertise, and us looking out for the discipline perspective, we should also keep a high bar to try to keep at least one or two within the field, for all the reasons I said.

CHAIR WALDROP: Okay. Cindi? More? Oh, sorry. Okay. Other thoughts, other comments? Are we ready to vote on the evidence? Jean-Luc?

MR. TILLY: So you're voting for evidence and here this is a health outcome. There are just two options, one for yes, and it passes, two for no, does not pass. I think we're missing just one more response in the room, if you all want to try again.

CHAIR MORRISON: Let's everybody --

CHAIR WALDROP: Can everyone try --

CHAIR MORRISON: -- re-vote one more time.

MEMBER BERMAN: As a reminder I'm a recusal.

MEMBER SANDERS: Wait, what was that last? Did you guys have to re-vote?

MS. ROILAND: No, that's fine, Amy. We're fine. Thank you, though.

MR. TILLY: Okay, great. Thanks to everyone for voting. The, we received 21 votes

in support of, yes, and one for no. So the measure passes on the evidence.

CHAIR WALDROP: Thank you. So we'll move on to considering gaps in care and opportunities for improvement, and also disparities. And I'll turn it back to Doug.

MEMBER NEE: Okay, terrific. So the description of performance data was presented over four years, from 2012 to 2015. Essentially the distribution of scores was reasonably consistent over that time period. And the range of means was between 61.4 and 66.4.

Standard deviation was a range of 20 to 24. Essentially, deciles of the facility scores in 2015 for 50 percentile. The score was 65 percent or less for this measure. And 65 percent or less of the hospices did score at this rate.

There really didn't seem to be, although there's no current established cutoff, but it really didn't seem to be any topping out of this measure as well. Disparity data was

looked at for race, ethnicity, gender, age, insurance status, and so forth.

Initial testing data included six months, a total of 1,409 patients. They looked at 463, or 30 percent, of patients who responded uncomfortable for pain.

The results showed that relative to cancer versus non-cancer pain, that the comfort level was 81 percent verse 84.8 percent respectively. No disparities there. No statistical significance was identified. But the percentages were pretty close.

No statistical significance was identified for ethnic distribution of patients whose pain was brought under control, versus not brought under control.

Further analysis in 2014 showed no statistical significance for a sample disparity relative to age, gender, or race. So the disparities appear to support the usability of the measure.

CHAIR WALDROP: Great. Thank you.

Anything to add, Debra? Okay. Comments or thoughts from the committee on gaps, opportunities for change? Okay. Seeing none I think we're ready to go ahead and vote on gap, and the performance gap. Jean-Luc.

MR. TILLY: This is for -- the polling is now open to vote on the performance gap. Your options are one high, two moderate, three low, and four insufficient.

Okay. So the polling is now closed. We received 16 votes for high and seven for moderate, zero for low, and zero for insufficient. So the measure passes the performance gap.

CHAIR WALDROP: Okay, thank you. We'll move on to reliability and validity. Debra or Doug?

MEMBER NEE: Yes. So, as far as reliability goes, just looking at numerator and denominator, exclusions as well. Data source is patient's self report of pain. There were really no issues or concerns regarding specifics in the

measure.

There was a little, there was a discussion actually on our workgroup call, relative to language barriers. And it was kind of thought that, maybe if there was an interpreter, or same interpreter asking the question, whether you had pain, and then asking the question afterwards, 48 hours later, if you achieved comfort, would be something that would be very helpful.

The initial testing, I'm going to, okay, so I'm going to qualify here. There was updated testing that was provided. And I'm going to let Dr. Scheffey talk about that.

But first, I thought I would just interject that the initial testing looked at 58 hospices and 38,000 patients, analyzed between 2009 and 2010.

The developer utilized inter-class correlation coefficient, and examined the agency level between, versus within variances of the measure, using the two years of data. Just

quickly looking at statistical significance that was set at .05.

People reporting being uncomfortable due to pain remain relatively constant. As expected, there were significant differences in the percent of hospices reporting patients uncomfortable due to pain on admission, verse within the hospice.

Indicated the ICC demonstrated over 75 percent consistency of results within hospices from quarter to quarter. There was a follow-up testing looking at, again, between and within hospice variation was found to be .71. A value of .7 is often regarded as minimum acceptable reliability value.

So that was initially the initial testing, and follow-up, relative to statistics. And if, there was a follow-up, updated testing done by NHPCO. But I thought I'd have Dr. Scheffey discuss that.

CHAIR WALDROP: Did you want to comment?

(Off microphone comment.)

MEMBER NEE: Sure. And that's great.

MR. SCHEFFEY: On the subject of reliability, I'll add a few comments about the data that's come in since the initial submission. We did get very much more data during the period 2012 through 2015.

I included in the, in this submission, information about stability, and some stability considerations. The stability is quite good. It -- that's the score -- that's, that addresses the question of score-level reliability.

On the individual-level reliability there was -- we stand on the data that was provided with the initial submission, which was an experiment on 236 patients. It was 96 percent agreement on how the patients reported their subjective impression of pain. I think that says enough right there. A top out of .91 is very good.

On the score level reliability, in spite of my comments about stabilities in the

scores, the NQF staff did complain about how, well, if he shows stability over time, that is not the whole story for reliability.

I point out, first, with this measure, there are limits to what you can do. But actually, we did mention a result that is effectively a split sample validation, a split sample of my ability.

I, this, NQF has accepted, sometimes, split sample work that's based on just random reassignment of patients within a hospice. And in our submission, in Section 2A2.3, we give a result about how close the score would come if we ran the test again and just supposed that we had a random new sample of patients.

That is, and we state, if the score was 58 percent to begin with, then we'd have an 80 percent chance of being between 48 percent and 68 percent.

That result, which we reported on the subject of reliability, is identical to the result you'd get if you were looking at a split

sample validation on a hospice with 100 patients in the denominator. So that is our evidence on reliability.

CHAIR WALDROP: Thank you very much for clarifying. Comments from the committee? Questions? Okay. Seeing none, I think we're ready to vote on reliability.

MR. TILLY: So, polling is now open to vote on reliability. Your options are, one high, two moderate, three low, and four insufficient. So the results are three for high, 18 for moderate, two for low, and zero for insufficient. So the measure passes reliability.

CHAIR WALDROP: Thank you. So we'll move on to validity. Will that be --

MEMBER NEE: Okay, so for validity, the developers compared response rates from two different wordings, comfortable level and acceptable level, with a follow-up question related to pain.

No new testing was, or testing data was not provided. However, it appears new

statistical data on the initial testing was provided in support of validity for the measure.

The initial test measure of score level included 212 of 686 patients from nine different hospice agencies. In response, 66 responded comfortable, 64 percent responded acceptable.

And I'll back up just for a second. So, when they were asked the question, they're asked these questions sequentially I'm presuming, correct? Yes, as I read it. Was your pain brought to a comfortable level? Was your pain brought to an acceptable level?

Just looking at these two responses, 96 percent of patients gave the, a same answer in total, kind of indicating good concurrent criterion validity for the measure.

Updating tests, updated testing looked at kappa values. And a kappa value of .91 was actually reached. Confidence intervals were not necessarily reported. But that was the kappa level that was identified, indicated a high level

of agreement between the two responses.

As far as threats to validity, there really was no identifiable threats to validity, no risk for the measures that was cited. Standard care for hospice is to provide timely and effective pain management, based on patient preference.

No exclusions. The patients were identified, are examined, measured. Exclusion criteria did not have any glaring, as far as I could see, groups inappropriately excluded from the measure.

Regarding meaningful differences in performance, they looked at 97 hospices, with greater than 50 patients in the denominator and found 16 per-- 16 hospices, or actually five percent, as statistically significant measure scores on the national average, with 292 hospices reporting.

I don't know, essentially, what that means relative to meaningful differences from a clinician's standpoint, looking at statistics

verse clinical effectiveness.

The developer indicated that samples used had very little missing data. However there was -- NQF identified that there was a failure to provide that data on frequency of missing data.

The only other thing that I had as a side note is it's -- one might argue, and it's not to deter anything here. But as far as acceptable and comfortable, sometimes we have patients that say their pain is acceptable, but they're really not comfortable. So I don't know if that really plays into here, but I just wanted to throw that in there.

NQF recommendations on this one were insufficient, I believe. Is that correct? Yes.

CHAIR WALDROP: Thank you. Any comments from the committee on sort of the concept of words matter? Is it acceptable? Is it comfortable? Any thoughts about that, or Doug's comments on missing this? Cindi?

MEMBER PURSLEY: I do want to come back to the perception from patients, and the way

we were able to improve our scores, because that's what we were looking at, we had to identify at admission, what is your most uncomfortable pain? Okay. So it's a knee. All right. And that's what we have to evaluate at 48 hours.

Because all of our patients have many other illnesses, many co-morbidities that result in discomfort. And so we had to narrow it down to one specific pain.

Because you can manage some and not others quite as well. And so, I think that that did create some issues. And that may be part of the differences in statistical outcomes.

CHAIR WALDROP: Okay. Thank you. Sean.

CHAIR MORRISON: I just had a question as to why there is no risk adjustment given the number of years this has been out there and the data points that you have. And why it hasn't been risk adjusted as an outcome measure.

MS. SPENCE: Well, on the basic

demographic information, there were no differences. So there was no basis there to risk adjust. What would you like to see it risk adjusted on?

CHAIR MORRISON: I'd like to see a couple of things. I think increasingly we are seeing too many differences in socioeconomic status. I think certainly we are seeing differences in -- regional differences. And so I have difficulty, and certainly within diagnosis, but the over and under 65 risk adjustment, without differences, doesn't seem to me --

MS. SPENCE: We did look at age.

CHAIR MORRISON: Yes. I said over and under 65.

MS. SPENCE: Yes, yes. Right.

CHAIR MORRISON: But the others I would certainly like to have seen, particularly, and perhaps co-morbidities.

CHAIR WALDROP: Cindi? Oh, other questions, comments? Cleanne?

MEMBER CASS: Yes. I have a little

concern about the comfort. Would it be better if the numerator were patients who had their pain reduced by two points or three points or something rather than bringing them down to completely comfortable? Is that safe and appropriate in 48 hours?

MS. SPENCE: So I understand your reasoning behind that. But it goes back to my point I was making when I was talking about it, that when you start putting clinician judgment in there or setting an arbitrary whatever, then you take that patient's voice out of it.

And the point of this measure was to have the patient define comfort for them. And when you start trying to then impose the assessment scales --

The other thing is that not everybody can use 0 to 10. You've got other's different assessment scales thrown in there: mild, moderate, severe; in some cases you're using faces. Because the clinician will adapt their assessment instrument to the needs and the

abilities of the patient.

CHAIR WALDROP: Yes.

MS. SPENCE: So because of that lack of uniformity of an assessment instrument across the board we chose to use comfort. We also did -- but I agree with you. It is a high bar. But we chose that on purpose.

The other thing, though, to keep in mind though, is that 100 percent should never, will never be the set threshold for this measure. There's types of pain that you cannot get under control within that 48 hour window.

You know, neuropathic pain is difficult. If the pain is existential in nature and being expressed physically, you're not going to be able to even figure that out and get it addressed in 48. So that is understood that that is not your goal.

And also, when you've got pain from multiple sources, you're also probably, and depending on the severity, depending on the long standing nature of it, you're probably not going to get

that all done in 48 hours.

So that part is maybe, perhaps wasn't explicit in the submission materials. But it certainly is part of the understanding behind the measure.

MEMBER CASS: Yes. And that would probably have to do with the acceptable percentile in the scores --

MS. SPENCE: Exactly.

MEMBER CASS: -- at some point. But as we're increasingly using methadone as well and bringing people into the inpatient unit, we know that they're going to get better over a few days. But it's going to be more gradual.

And our clinicians have a hard time with this criterion. They feel like they could put patients in jeopardy if they were expected to get them completely comfortable.

MS. SPENCE: And exactly -- yes. And so again, you know, there were no thresholds set. We didn't get to that point with it. But they absolutely would not be at 100 percent. Nor

should hospices.

And, I mean, and that would be made clear also, that we're not encouraging overdosing patients with pain medication with this measure.

CHAIR WALDROP: Thank you for the clarification. Amy.

MEMBER BERMAN: I apologize. I was recused from an ASCO measure, not your measure. So I'm participating. So, I have a question. What are the implications, for example, of a patient that might have been a drug abuser, somebody who is not medication naive?

Are there exceptions, exclusions to this kind of a measure? Are there things that you have to consider?

And the other comment kind of goes to Cindi's comment. And you're making such brilliant comments. But as a human being on the receiving end of care, I don't want somebody to look at a body part; I want them to look at me overall.

The notion of pain is kind of an

absolute. Am I uncomfortable, or am I comfortable? And this distinction in whether or not the pain has been resolved, the acceptable component is I think closer to what is likely true.

There are tradeoffs. And there may be tradeoffs. But is it an acceptable level for that patient is probably closer to right from a human being perspective.

MS. SPENCE: So, Amy, to your point, both of your points of that patient who says, yes, I'm uncomfortable, but I'm going to refuse pain medication because of my history with addiction or whatever reason.

And also, that patient who, maybe they have multiple sources of pain, making it complex to treat them, that also goes back to where you would set the threshold.

So in neither of those cases would the expectation be that that patient be brought to a comfortable level within the 48 hours.

So those patients, they wouldn't be

considered exceptions because we're not exempting them from the measure. But there's where you go to, looking at that.

Perhaps, Sean, even risk adjusting on the proportion of those patients. Risk adjusting on things like short length of stay for those patients that can't re-self report at that second follow-up, which is also an issue. And that came out. Both of those things came out in the reporting.

So just one note of clarification. The measure itself, the question is, are you comfortable? The acceptable word was put in the validity testing, with the idea that while the two --- and I agree with you, Doug, that they're not necessarily comparable. But if the -- what we were looking at was, would the patient's response change if that -- with that word acceptable? And the answer came back, no. So therefore, they showed that they were comparable.

We had questions going in. Would they turn out to be comparable? They did turn out to

be that. But semantically, yes, we recognize there's a difference there.

And we went with comfortable, again, with the idea of, if that set the bar higher, that's what we wanted. But it turned out that it did not with the testing.

CHAIR WALDROP: So, I'm going to thank you for your clarification, move -- take one more comment, and move us to a vote to keep us on track with time here. So, Cindi.

MEMBER PURSLEY: To Amy's point, and also the developer's. If questions like these are yes or no, it doesn't take in for the variations that the vast majority of hospice patients, that's where they lie.

They're not going to be, you know, if you ask most hospice patients, are you uncomfortable because of pain? Yes. Was your pain brought to a tolerable level within 48 hours? Again, it comes back to, if it's no, that is a score that that hospice is held to.

And we want to be able to show that we

are doing the very best for our patients. And those questions don't always allow that to shine through.

CHAIR WALDROP: Thank you. I'm going to move us towards voting on validity, but first I just want to ask Karen to say brief word about the insufficient assessment by the staff.

MS. JOHNSON: So, why did we land on insufficient? Really just a couple of things. And one I don't think we've hit. Although, I believe we talked about it in the workgroup call.

You didn't talk about exclusions. And I think it might be just a miscommunication in the submission about what is an exclusion. Maybe some people aren't -- maybe you can just speak to that in a minute.

The other thing that we noticed is, we were curious, as Sean asked, about the risk adjustment. And you do have some information about not seeing differences in those factors.

Our question was, did you report patient level data? Or did you report agency-

level data. Because if it's uniform across agencies, then we would agree that risk adjustment class probably isn't needed. But it looked like it was patient-level data that was shown. So, I think that was the question there.

And that doesn't necessarily address all the things that you could have looked at. And we didn't get into that. And let's see, I'm looking, is there anything else?

The final comment, additional score level validity testing may also be needed. That's only if you don't really accept that doing the comfortable versus acceptable is a valid demonstration of validity.

To be honest with you, we had two or three of us look at it internally. And one said, yes, it was. The other said I don't think it is. So we put it out there for you. This is for you to decide if that is an okay check to validity.

CHAIR WALDROP: Thank you for that. Did you want to comment?

MS. SPENCE: First, let me just start

by saying I think we -- this should be there somewhere in the submission, but just to clarify. Between 46 to 48 percent of patients say they have no pain on admission. So it's not the majority who are saying that they have pain.

I'm going to ask -- well, in terms of the exclusions, you're right. On our call it came up. It's a little bit of semantics going on here. Again, as the measure is in two parts, our exclusions are sort of in two parts. You can call the first one defining the sample, if you want to.

So the people who are not considered for the measure are those who are under 18, where there's no way for them to understand the question because of a language problem, and using both interpreters, including family interpreters, is fine. And this came out in the implementation when CMS had it implemented. We made that clear.

And then the patients who cannot self report, and either they can't self report because of their underlying disease makes them

cognitively unable to report, or they're so ill that the cannot self report.

So that group is not considered for the measure to start with. In a purer sense, the exclusion are those people who say they're not comfortable, I mean they are comfortable. So therefore they're not uncomfortable. Therefore, that's the true exclusion group for the measure. And we did look at those.

CHAIR WALDROP: Thank you.

MS. SPENCE: And was there one more, Karen, about the risk?

MS. JOHNSON: Yes, the risk factors. Was that patient-level data that you showed, or was that agency-level data?

MR. SCHEFFEY: Okay, sorry. The data we reported was patient-level. We have a little bit of facility-level data we could use for that. We could pursue it more in the future if we got more data in, had more power, yes.

CHAIR WALDROP: Okay, I think we need to move towards voting on validity of Measure

0209.

MEMBER STEINBERG: Just one quick question. I'd just be curious to know what the discussants and you guys on the work group thought of it. I mean, to me it seems more kind of like moderate. So I'm just -- anything?

MEMBER NEE: We really didn't have a whole lot to discuss on the work group call relevant to comfortable, acceptable, that sort of thing. For what appears to be a really simple measure, there sure is a whole of passion, and concern, and angst associated with it.

So just from a clinician standpoint, looking at this, it really is tough. And what other words would you look at? And looking at the exclusion data we had, as Carol had mentioned, we did have some conversation about that, and what could or couldn't, and what constitutes -- you know, voting age of 18 versus how low do you go in age as far as things like that occur. But, yes, we didn't have too much more than that.

CHAIR WALDROP: Thank you. And I believe Alice on the phone has a question for us. Alice?

MEMBER LIND: I sort of lost track of when you were going through the exclusions. Did you mean that people who use family members to translate for them are included as long as you get a translated answer to the question?

CHAIR WALDROP: I'm getting a yes from Carol. You can't see her nod her head, but I can. Okay. I think that brings us to a vote on validity. Jean-Luc?

MR. TILLY: So, yes, polling is open to vote on validity. Your options are one: high; two: moderate; three for low; and four for insufficient.

CHAIR WALDROP: So perhaps you can interpret for us.

MR. TILLY: Yes. So the results are one high, eight moderate, six low, and eight insufficient, which is a consensus not reached.

MS. JOHNSON: Let's do our math here.

We've got one plus eight is nine out of 23.

MR. TILLY: Right.

MS. JOHNSON: We want to be careful. I'm seeing that as 39.

MR. TILLY: Yes, you're right. Yes --

-

MS. JOHNSON: So at this point, it actually did not pass validity. Because it has to be 40 to 60 inclusive to get that.

I think what we can do, because there was a lot of insufficient. A lot of people felt that there wasn't enough information here. Let's find out what you'd like to see. This is not the end of the game, necessarily.

We will be having a post-comment call. And we could potentially look at some extra information if it's something that you could pull for us, Carol.

I'm thinking one of the things might be the agency-level age distribution to indicate risk adjustment potential. I'm not sure if there are others. Maybe we can just have the committee

first tell us why you landed on insufficient. What would you like to see to get you out of insufficient? We'll find out if that's possible for Carol to do.

CHAIR WALDROP: Christine?

MEMBER RITCHIE: I think more data around or a plan to risk adjustment.

CHAIR WALDROP: Woody?

MEMBER MOSS: I was concerned about the 16 out of the 97 hospices where we didn't have a very good feel on what was going on and why they fell out of the target. So I had concerns about the data there.

CHAIR WALDROP: Other suggestions for how to make the data less -- or more sufficient?

MS. JOHNSON: Do you think this is something you have about six weeks or so before this next fall is ---

MS. SPENCE: So the thing is is that when we do our national data collection, we collect aggregated data. The patient-level data that we had, we did as a concerted effort to ask

for patient-level data from several hospices so we would have that. Because it's not part of or normal data collection.

I don't that in -- so age is not something, for example, that we had it from those hospices, which is why we did it at patient-level. We don't have a lot of hospices that we have that level of data.

So I don't that within six we'd be able to get it. We can try. But I can't guarantee that we'd be able to.

So I just want to be clear. Is the focus here mostly risk assessment? Because I want to able for us to focus our --- I could certainly come up with a plan for risk assessment.

Because I have ideas for risk assessment that go beyond your regular just demographic, that I think are really more germane to this measure in terms of digging down and really showing where the meaningful differences are. If a plan -- I can certainly present a

plan.

CHAIR WALDROP: Sean, then Paul, then Amy.

CHAIR MORRISON: Yes. I'm not sure. I think what I'm hearing, Carol, when I hear it, is that there's, and the committee can correct me, I think there are some concerns about the risk adjustment. I'm not sure that people are specifically asking for a risk adjusted model. But I think at least some interagency data would be helpful. And I'm not sure that would be too difficult for you guys to provide. Is that fair?

MR. SCHEFFEY: Well, the trouble with that is many of our hospices are very mixed in their populations. It's not as though I can say, ah, here I have five hospices that have purely urban populations and low SES.

MS. SPENCE: But we can, I mean, we can look at it, Sean, and see what we can do.

(Off microphone comment.)

MS. SPENCE: Yes, right. So we can at least tell you what we were able to do.

CHAIR WALDROP: Paul?

MEMBER TATUM: I was just going to argue for at least regional adjustments.

CHAIR WALDROP: Yes, Amy.

MEMBER BERMAN: Just a process clarification. So when I see these numbers here, so 9 would have been 4 out of 15 if we were to look at the insufficient as a separate, whether we wanted an exception or non-exception. Am I misunderstanding?

MS. SPENCE: Just a little bit. Exception only counts for evidence. So we don't have an exception for validity like we did for evidence. So basically, what we're looking at is high and moderate, is that enough to get you to at least that 40 percent mark or not? And the math got us at 39 percent.

But we'll talk with Carol offline. We'll see what's possible. Again, we can revisit at post-comment. It might be that you have the data, and you can allay fears of, or fears is the wrong word, but concerns about risk adjustment.

And that would be fair.

CHAIR WALDROP: Process clarification, do we vote on feasibility, or do we stop?

MS. JOHNSON: Let's go ahead and stop. We are a little bit along the way in our time. So let's stop. If they bring stuff back to us at post-comment, and it makes it past the validity hurdle, at that point we'll talk about feasibility and usability and use.

CHAIR WALDROP: So that concludes our discussion of 0209. Just being process-oriented, it's 12:25. We have stopped with three more measures by now. Do we just keep going?

(No response.)

CHAIR WALDROP: Okay. So we'll keep on going. We'll move on to Measure 1634.

MEMBER MOSS: Deborah?

CHAIR WALDROP: Yes? Sorry, Woody.

MEMBER MOSS: So I'm just wondering about process. At this rate, we're going to be here until sometime next week. Sean brought that up in his introductory remarks.

I was thinking we would have supposedly scanned these things at least in advance. And we should be pretty snappy in our responses and questions and moving right through them. Karen, is that unfounded, or is that what most of your groups do?

MS. JOHNSON: It depends. My favorite answer. Some things go a lot faster than others. Some groups never go fast. Others go really fast. We laugh about the surgeons, how fast they go through things. And the endocrinologists don't. So it kind of depends.

CHAIR WALDROP: So maybe we can try just stepping it up a bit and seeing if we can get through next couple in a shorter period of time.

CHAIR MORRISON: As Laura brings up, I think, Woody, you raised a good point. I think there are a couple of things that help. One is to be ahead of and be prepared with your comments about what you want to run through them.

The second is really to speak to data

and not anecdote. And the third, as we talked about at the beginning, people can get very emotional about these. And if a point's been made, it doesn't need to be repeated. Otherwise, I mean, Doug and are completely comfortable being here until as long as it takes. We like hanging out with you guys.

CHAIR WALDROP: Okay. So I'd like to welcome Laura Hanson back and ask you to give us a very brief overview of 1634, pain screening.

MS. HANSON: I'm so terrified. I'm between you and lunch. But very briefly, the pain screening piece measure is very similar to the Dyspnea screening measure that you recently reviewed and is basically also submitted as a paired measure.

So it's paired with pain assessment in that the results from a pain screening measure define the denominator population for individuals who should have a pain assessment.

Also, similar to what you've heard before for this measure, we're submitting data to

remove the exclusion for hospice -- individuals enrolled in hospice who have been enrolled less than seven days. And I'll stop there.

CHAIR WALDROP: Thank you very much. So I'll move it to our discussants who are Bob Archuleta and Michelle Caughey. And if you could please give us an overview of the evidence.

MEMBER CAUGHEY: I would like to say I think we can go fairly quickly on these two measures. They are paired. I wanted to thank the committee, the subcommittee. They were very, very helpful. And you've been exceedingly helpful on discussion. I'd like to move to the punch line, which is insufficient evidence with a recommendation for evidence -- or exception.

CHAIR WALDROP: So any more about that suggestion?

MEMBER CAUGHEY: Well, I think that, you know, there's general agreement by every clinician that assessing for pain is extremely important.

CHAIR WALDROP: Okay. Comments from

the committee, thoughts about this?

MS. JOHNSON: So just to be sure, this would be like the other one. If you feel like there is insufficient evidence linking screening to better outcomes, then you would vote insufficient as we did before. If more than 60 percent of you vote insufficient, then we have the option of thinking about going the exception route.

CHAIR WALDROP: Christine?

MEMBER RITCHIE: This is a quick reminder that the paper we've mentioned looked at both pain and dyspnea. Thank you.

CHAIR WALDROP: Any other comments, thoughts about this measure? Seeing none, I believe we are ready to vote on the evidence. Just want to remind you that if we vote insufficient then we will be -- if a significant number of us vote insufficient, then there will be the following vote to consider an exception.

MR. TILLY: That's right, yes. So the polling is now open evidence. Please vote 1 for

high, 2 for moderate, 3 for low, and 4 for insufficient.

The results are 2 for high, 2 for moderate, 0 for low, and 19 for insufficient. So the measure goes insufficient.

So the vote is now open for evidence of -- a potential exception to critical evidence. Here we have just two options. 1 for insufficient evidence with exception, and 2 for no exception.

So the vote is unanimous. 23 voted 1, insufficient exception --- insufficient evidence with exception. And none voted for no exception.

CHAIR WALDROP: Okay, thank you. So I'd like to move on now to look at gaps and opportunities for improvement. Michelle?

MEMBER CAUGHEY: Yes, thank you. There are clearly opportunities for improvement among hospices in doing this, which surprised me. So there is definitely a moderate opportunity.

There were, interestingly, no disparities identified in the 1b measure. So I

think the preliminary rating of moderate is appropriate. I think there are opportunities for improvement.

I will mention and remind you that we have good data for hospice and that we will be having data for palliative care. So I'm only reporting palliative care because we've already had that discussion, unless others want to have that discussion.

CHAIR WALDROP: Okay, Christine? Did you have a question? I'm sorry, okay. I just had to make sure. Comments or questions on gaps and opportunities for improvement with this measure? Now we're rolling here. Okay. So I'd like to move us then to consider a vote for gaps and opportunities for improvement.

MR. TILLY: The polling is now open for the performance gap. Please vote 1 for high, 2 for moderate, 3 for low, and 4 for insufficient.

The results are 1 for high, 19 for moderate, 2 for low, 1 for insufficient. The

measure passes for performance gap.

CHAIR WALDROP: Thank you. So we'll move on to consider reliability and the specifications. Michelle?

MEMBER CAUGHEY: For reliability, I agree with the assessment of moderate reliability. You'll recall that, again, for the hospice item set that there are a large number of hospices reporting with very large numbers of patients. And so it is a reliable measure in the hospice item set.

And inter-rater reliability is also moderately high. For palliative care, there's some information about inter-rater reliability as well but in a small number of patients.

CHAIR WALDROP: Okay. Should I just ask if Bob is on the line? Anything else to add, Bob?

MEMBER ARCHULETA: No, other than I think the percentage of reporting by hospices was in the range of 82 to 100 percent. So obviously it was very high reliability.

CHAIR WALDROP: Great. Thank you. Any comments or questions or thoughts from the committee? Okay. Seeing none, I'd like to move us to a vote on reliability.

MR. TILLY: Yes. The poll is now open for reliability. Please vote 1 for high, 2 for moderate, 3 for low, and 4 for insufficient.

The results are 5 voting for high, 18 voting for moderate, 0 voting for low, and 0 voting for insufficient. The measure passes on reliability.

CHAIR WALDROP: Okay, thank you. So that moves us into the area of validity, validity testing, and any threats to validity. And I'll go back to Michelle and/or Bob, whoever was going to take this position.

MEMBER CAUGHEY: Bob knows more about statistics than I do. Please chime in.

MEMBER ARCHULETA: I'm good with your discussing it, since ---

(Laughter.)

MEMBER ARCHULETA: -- you know the

committee's take.

MEMBER CAUGHEY: Well, definitely our subcommittee agreed with the rating for moderate validity.

So there was some missing data, but it was very, very small. What else should I say about validity, which I don't completely understand from a statistical point of view? I never did, and I'll ask Karen if you have anything to add.

MS. JOHNSON: Again, just when you're thinking about testing specifically, you think about the scope of the sample that was used in the testing. You think about the method. The methods were appropriate. And then you look at the results. And we did try to put the results and develop our thinking as well, try to put those into some kind of context for you.

CHAIR WALDROP: Thanks. Any threats to the validity that you saw of significance?

MEMBER CAUGHEY: Again, only in the palliative care realm and not in the hospice

realm.

CHAIR WALDROP: Okay. Thanks. Seeing no further comments from the committee, I'd like to move us to a vote for validity on 1634.

MR. TILLY: And the polling is now open for validity. Please vote 1 for high, 2 for moderate, 3 for low, and 4 for insufficient.

The results are 1 voting high, 22 moderate, 0 for low, and 0 for insufficient. So the measure passes validity.

CHAIR WALDROP: Okay. So I'm learning that we can consider feasibility and usability in use through the same lens that we considered the dyspnea measure some hours ago. We've still got to open --- I want to open that just to see if there are any objections, since the methods are the same. I'm wondering if there's any objections to using the vote that we captured earlier. Any problems with it?

(No response.)

CHAIR WALDROP: Okay. Then we will cut and paste that vote and move on to the next -

- to the overall measure, the overall vote of acceptability of this measure. So when you're ready, Jean-Luc, let us know.

MEMBER RITCHIE: Is it possible to review what we just said? Can you review for us, please, just review the votes?

MR. TILLY: Yes. So on evidence, we voted insufficient with exception. Oh, sure, sure.

MS. JOHNSON: Let's clarify, Christine. Did you want to know the votes for this measure so that we can do the overall? Okay. So carry on, Jean.

MR. TILLY: Very good. All right. So, yes, we voted first insufficient, and then we passed it with insufficient evidence with exception. And we passed it on performance gap, 1 high, 19 moderate, 2 low, 1 insufficient. On reliability, 5 high, 18 moderate.

CHAIR WALDROP: Okay. So let's move to the overall vote of acceptability of this measure.

MR. TILLY: Okay, the polling is now open for overall suitability for endorsement. Please select 1 for yes or 2 for no.

The vote is unanimous. 23 vote yes; 0 vote no. The measure passes.

CHAIR WALDROP: So the measure passes. Thank you. So we'll move on to the assessment of -- the pain assessment measure, 1637. I'd like to begin by asking Laura if she has anything further she wants to add about this particular measure.

MS. HANSON: Just very briefly that this measure draws its denominator from the one that you just voted on. So for individuals who expressed pain at the time of initial comprehensive assessment in hospice or in palliative care, then the numerator condition here is to include at least five of seven elements of a comprehensive pain assessment in order to pass this quality measure. And likewise, this is one where we're looking for removal of the seven day exception.

CHAIR WALDROP: Thank you. And we'll stay with our discussants, Bob and Michelle. So I'll go back to you, Michelle.

MEMBER CAUGHEY: This is 1637. And the votes -- or what I would recommend is parallel to what we just went through which is the evidence is insufficient for the same reason. But the committee recommended, or we recommended an exception to that, that the quality of pain as described by the patient is extremely important to the clinician and to the patient, and that we recommend it be kept as a measure.

CHAIR WALDROP: Okay. Any comments or thoughts or concerns from the committee?

MS. JOHNSON: I have a note to myself that we had some additional evidence provided to us after the work group call. Is that what you were going to remind us of? So maybe you can just briefly describe what that was.

MS. HANSON: So just briefly, because there was concern about the evidence link with outcomes, I should give Kathryn Wessell the

credit, my colleague, my esteemed colleague who helps and works so vigorously on this project.

We basically looked for additional information, and I believe the attachment was offered to you, and basically found an additional systematic review as well as a couple of other individual publications that link systematic pain assessment, specifically with some patient-level outcomes: shorter length of stay in the ICU, less time on mechanical ventilation, decreased pain intensity, and decreases in adverse health events. So a modest body of evidence, but additional to what we had submitted.

CHAIR WALDROP: Thank you. And that was short turnaround. So we appreciate your supplying that additional information. Does that raise any questions or comments from the committee that we need to hear? Looks like we have one here.

MEMBER MOSS: So, yes, so is that tangential evidence, or is that just a little bit more evidence, but it's not enough to persuade us

that there's insufficient evidence?

MS. HANSON: It depends. I think that's the answer. I think what is actually important about the additional evidence that we submitted, although it is limited primarily to the critical illness population, it is an interesting link between pretty compelling patient outcomes and assessing pain in a routine manner. So it depends.

CHAIR WALDROP: Other -- Michelle.

MEMBER CAUGHEY: Well, thank you. But I think, you know, it was in a limited setting. And we're recommending a measure across many settings. And so I think that is interesting. But that didn't change my thinking.

CHAIR WALDROP: Okay. Seeing no other comments, I believe we move towards vote on the evidence for measure 1637.

MR. TILLY: So for evidence, your options are 1: high, 2: moderate, 3: low, and 4: insufficient.

Okay. And the results were 1 high, 9

voting moderate, 0 voting low, and 13 voting insufficient. So I believe the measure does not pass.

MS. JOHNSON: You are correct. We have to have more than 60 percent on insufficient for this to go through. And the percentages came out to 56.5 percent. So in this case, I think the additional evidence kind of split our vote a little bit there. So right now, it has died on evidence. Yes, it did not pass evidence.

PARTICIPANT: And no exception?

MS. JOHNSON: We could not get to exception.

MEMBER STEINBERG: Is it ever possible to re-vote?

MEMBER HANDZO: Karen, I didn't hear the end of your last comment. It kind of trailed off.

MS. JOHNSON: Oh, sorry about that. Yes, it did not pass evidence with this vote.

CHAIR WALDROP: Can I just ask, I think there is some confusion about whether the

exception is just for evidence or all others. And I'm wondering, to pick up on your point, if we could, if it's possible to have a re-vote for people with the clear stipulation that only with evidence can we get to an exception. Is that correct?

MS. JOHNSON: Right. So what you have to ask yourself with the additional evidence that was provided is is that enough to get you over the hump. And by enough, it has to be -- you have to feel like that's the full body of evidence that would link in your mind assessment of pain to outcomes.

CHAIR WALDROP: Okay.

MS. JOHNSON: Again, we have to have more than 60 percent to get to insufficient. If we get there, then you can go on to potential exception.

CHAIR WALDROP: It's only possible to do this with the evidence vote.

MS. JOHNSON: Yes, right. So this is not something that's possible for reliability or

validity.

CHAIR WALDROP: Yes.

MEMBER MOSS: So can I read the first sentence of what Laura has sent us that's the updated evidence summary, if that's all right?

The first sentence says, "While there were no new high quality systematic reviews linking pain assessment to improved patient outcomes, we were able to find some additional evidence to support the measure."

And then when you go down and look, some of them are not the world's best class journals, BioMed Research International and things of that nature. That's the reason why I voted insufficient, okay. But I think it did -- this discussion was confusing.

MEMBER STEINBERG: So I don't know if it's required, but I'd like to make a motion for a re-vote.

CHAIR WALDROP: Yes. And I think, it sounds like it's possible. So any objections to taking a re-vote, another vote?

(No audible response)

CHAIR WALDROP: Okay. So let us know when we're ready on the --

MEMBER TATUM: My understanding is we're one vote off being able to move forward in the insufficient category. Is that correct?

CHAIR WALDROP: Yes.

MEMBER TATUM: Thank you.

CHAIR WALDROP: Did you have a question, Christine? Did you have a question? Your card, sorry.

MR. TILLY: Okay, the polling is now open for a second vote on evidence. Your options are 1: high, 2: moderate, 3: low, and 4: insufficient. The votes are now 0 for high, 0 for moderate, 0 for low, and 23 for insufficient. So the measure vote is insufficient.

The polling is now open to vote on whether the measure is insufficient evidence with an exception or whether no exception. The votes are 24 for insufficient evidence with exception and 0 for no exception. So the measure passes

evidence with exception.

CHAIR WALDROP: Okay, so we'll move on to considering gaps and opportunities for improvement. And I'll go back to Michelle please.

MEMBER CAUGHEY: This measure was difficult for many hospices to do because imbedded in it are several different -- the characterization of the pain, so they have to meet five elements out of six. Am I right about that? Oh, out of seven. And so there was great opportunity to improve at that level, but no disparities were seen.

CHAIR WALDROP: Okay. It's open to the floor. Any comments, thoughts, questions about gaps and opportunities for care, opportunities for improvement rather? Seeing none, we can move towards voting on opportunities for improvement.

MR. TILLY: The vote is now open on performance gap. Select 1 for high, 2 for moderate, 3 for low, and 4 for insufficient.

MS. ROILAND: Alice, can you submit your vote for performance gap? The options are 1 high, 2 moderate, 3 low, and 4 insufficient.

MR. TILLY: The results are 11 voting high, 13 voting moderate, 0 voting low, 0 voting insufficient. The measure passes performance gap.

CHAIR WALDROP: Great. Thank you. So we'll move on to consider the scientific acceptability of the measure properties. We'll begin with reliability and reliability testing. Bob and Michelle?

MEMBER CAUGHEY: Very similar to the other measures at moderate reliability, inter-rater reliability was fairly high both for -- it was high for palliative care and moderately high for hospice.

CHAIR WALDROP: Okay, thank you. Any comments from the Committee, questions or concerns? I want to make sure Bob has a chance to --

MEMBER CAUGHEY: Yes, I should say

Bob, anything that you would like to add?

MEMBER ARCHULETA: No, I think that's -- what was said was the important part.

CHAIR WALDROP: I would like to ask Karen if you could clarify for us the preliminary rating as moderate and insufficient. Is that a typo on the measure evaluation script?

MS. JOHNSON: I think we have a little bit of a disconnect because the one that we're showing here --

CHAIR WALDROP: Is moderate.

MS. JOHNSON: -- is not checked.

CHAIR WALDROP: Okay.

MS. JOHNSON: So it may have been insufficient first, and you somehow or another may have an old version, but I'm not quite sure.

CHAIR WALDROP: I just wanted to clarify, thank you. So I believe we're ready to vote on reliability when Jean-Luc lets us in.

MR. TILLY: Yes, the polling is now open to vote on reliability. Select 1 for high, 2 for moderate, 3 for low, 4 for insufficient.

The results are 3 voting high, 20 voting moderate, 0 voting low, and 1 voting insufficient. The measure passes reliability.

CHAIR WALDROP: Okay, we'll move on to validity, validity testing, and any threats to validity. And we'll go back to Michelle and Bob.

MEMBER CAUGHEY: So again, moderate as with the other measure. The only threats to validity had to do with the palliative side versus the hospice.

CHAIR WALDROP: Anything else to add, Bob?

MEMBER ARCHULETA: No, just that the statistics for palliative care to show the validity should be forthcoming, but it just isn't quite there yet.

CHAIR WALDROP: Okay, thanks. We open it to the committee. Any comments or thoughts or concerns, or questions from committee members? Seeing none, I think we're ready to vote on the validity of Measure 1637.

MR. TILLY: All right, to vote for

validity, select 1 for high, 2 for moderate, 3 for low, and 4 for insufficient. Okay, the results are 0 voting high, 24 voting moderate, 0 voting low, and 0 voting insufficient. The measure passes validity.

CHAIR WALDROP: Thank you. So we'll move on to feasibility or the ease of data collection and capture. And I'll go back to Michelle please.

MEMBER CAUGHEY: Can we do that other process that we did before? Yes? Okay.

CHAIR WALDROP: Thank you for reminding me, yes. I'll get this. So I want to remind you we can go back to our vote from earlier this morning when we looked at the dyspnea measure where we considered feasibility and usability and use and we can consider the same vote if we want to carry that forward.

Do I see any objections to carrying our vote forward? Seeing none, we can move on then to the overall vote of acceptability of this measure.

MR. TILLY: Polling is now open to vote for overall suitability for endorsement. Select 1 for yes and 2 for no.

The results are 24 voting yes and 0 voting no. The measure is recommended.

CHAIR WALDROP: Okay. I believe that we have one measure left in this session, but we are going to stop for now and take a break. So I just want to thank you all very much for indulging me in this new role. Thank you so much for making this easier, and I look forward to talking with you over lunch, and we'll come back together afterwards.

CHAIR MORRISON: Actually, we're not going to take a break yet. Sorry, sorry, sorry. First we're going to do public comment on the morning session. So could I ask, Operator, could you open up the phones for public comment for people on the telephone lines?

OPERATOR: Yes, sir. At this time, if you would like to make a comment, please press star and then the number 1.

CHAIR MORRISON: And could I just ask you to state your name, where you're from, and your public comment succinctly.

OPERATOR: And at this time there are no public comments from the phone line.

CHAIR MORRISON: And then I would like to invite anybody in the back of the room, if you would like to make a comment, please step up to the microphone.

MS. AST: Hi, everyone. I'm Katherine Ast, the Director of Quality and Research for the American Academy of Hospice and Palliative Medicine.

AAHPM and other organizations from the National Coalition of Hospice and Palliative Care are here to express our strong support for the continued endorsement of all the measures brought forward for maintenance in this project.

Please take note of the letter we submitted prior to this meeting which highlights some of the issues our field faces that contribute to our lack of relevant measures,

particularly those with a true palliative care denominator.

What we want to emphasize today is how critical it is that we keep the endorsement of the measures we do have so they can be used to improve the quality of care for our patients and families and to enable our clinicians to participate in value-based reimbursement.

NQF, CMS, and the MAP have indicated through various publications and rulemaking that palliative care and end of life care represents a major gap in quality measurement.

Our field is very unique: since our patients are seriously ill, death is not always a negative outcome, can likely be a neutral or positive outcome.

We need measures that are flexible, take our uniqueness into account, emphasize care coordination, family meetings, goals of care, et cetera. The approach to measure development for our field cannot be cookie cutter.

In order to increase the usability of

the measures we have and expand the settings and populations for which they can be implemented, we need to keep working with what we have and keep the measures we do have endorsed.

We have so few outcome measures in our field, particularly patient-reported measures, and for good reasons. It's difficult to obtain the patient report in this population.

However, we do have NQF number 0209, the comfortable dying measure which is able to capture patient self-report of pain. Karen Johnson asked the question at the beginning of today's meeting that if we had outcome measures to capture enough aspects of the quality of care for patients with serious illness, would we still need process measures?

Unfortunately we have so few outcome measures that we couldn't possibly dispose of our process measures to measure quality. However we do have this outcome measure, and we believe we should keep it.

We believe that risk adjustment or

risk stratification is not critical for this or any other measures brought forth today although it is currently being explored for several of the measures.

The reason is the measures are used for comparison among similar providers, and there's no expectation that the performance be 100 percent.

Benchmarking is a critical component to measuring the quality of care. And without measures to report and data to aggregate, we can never get to any benchmarks in our field.

There are certain processes that many believe should continue to be measured, even up to a rate of 100 percent. What does it mean for a measure to be topped out?

With such a new field finding its place in healthcare and in different settings, we think all the measures are far from being topped out even if they approach 100 percent performance.

In addition, many measures continue to

show a clear opportunity for improvement. Once we expand the measures to be reported in multiple settings and with a true palliative care denominator, then we can start to enable benchmarking and true comparison of providers.

We'll need to keep the endorsement of all the measures presented here today in order to see that goal become a reality. Thank you very much.

CHAIR MORRISON: Other comments from the back of the room? So I think everybody has earned a break for lunch. However, I am going to suggest that that be a very short ten minute grab lunch, use the restroom if you need it, and let's continue with a working lunch for the afternoon if that works for people. Thank you for a very great morning. Hard work.

(Whereupon, the above-entitled matter went off the record at 1:05 p.m. and resumed at 1:26 p.m.)

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

1:26 p.m.

CHAIR WALDROP: Okay. Welcome back, everyone. We are going to finish this round by considering Measure 1628, Patients with advanced cancer screened for pain in outpatient visits. I will turn to our measure developer, who is joining us by phone, first, and then I will go to our discussants. We have two discussants present with us, Karl and Tracy, and then we have Amy Sanders on the phone. So I would like to begin by asking our measure developer, are you there and would you like to share briefly a little bit about this measure?

MR. LORENZ: This is Karl Lorenz. I don't know if I'm the only one who's present, is that the case? I don't know if Neil is on the line.

MR. WENGER: No, I'm here too, but for this Measure, let me let you talk, Karl.

MR. LORENZ: Okay. So this -- I think the measure we're looking at currently is routine

screening for pain. And this comes from the Cancer Quality ASSIST Indicator Set and in general it relies on evidence related to the use of screening for pain in every intervention related to either critical intervention or quality improvement in the field of pain research. And although -- I just would stress as well that in general adherence to this tends to be moderately high, but evidence looking at facility level variations still shows important sub-populations of patients and of facilities that are non-adherent to it, even in the Department of Veterans Affairs, in previous published works. But in general, it is intended to emphasize the need to screen for pain routinely in advanced illness.

CHAIR WALDROP: Thank you very much. So I would like to move to our discussants and ask, I believe Karl is going to start us off looking at the importance of this measure and of reporting it.

MEMBER STEINBERG: Great. Well, thank

you and I'm going to do the best I can with my eyes and if it doesn't work out, I may need to ask for help from my co-discussant. So this is somewhat similar to 1634 that we already discussed, except that this is a measure in the outpatient setting and this is for people with advanced cancer. It's a process measure. And essentially, they have to be screened for pain with a standardized tool at every outpatient visit. It's a maintenance measure and it's -- essentially, the evidence, there's some systematic reviews, there's not a lot of empirical evidence, but the idea behind it being just like the other measures that in order to treat pain, you have to assess whether they're having pain or not. And last time, it was Insufficient, but with an exception. I don't know if there's any specific questions about it.

CHAIR WALDROP: Okay. So let me ask Amy Sanders and Tracy Schroepfer if you have anything else you want to add about the evidence to this measure?

MEMBER SCHROEPFER: This is Tracy. I don't.

CHAIR WALDROP: Amy, are you there?

MEMBER SANDERS: It took me a second to find you to unmute you. I don't have anything to add in particular, other than that the Developer stated that there was no new evidence since this measure was initially endorsed in 2012.

CHAIR WALDROP: Thank --

MEMBER STEINBERG: Sorry --

CHAIR WALDROP: Okay.

MEMBER STEINBERG: -- one other thing. We actually did get some additional evidence that was submitted and it's in the packet now. And probably has to do more with kind of gap analysis and so on. But what was interesting was that this was a VA study from 2014 from Oishi and it indicates that basically the mean on this measure was 98 percent in the VA clinics. So the developers say that there are still some gaps, and whether it's topped out or not, it's just something that's probably worth mentioning.

CHAIR WALDROP: Thank you. So, did you want to comment on the Staff's recommendation of Insufficient with an Exception? Any thoughts about that or --

MEMBER STEINBERG: That seems reasonable.

CHAIR WALDROP: Okay. Comments from --

MEMBER SANDERS: Agree.

CHAIR WALDROP: Oh, Amy, sorry.

MEMBER SANDERS: I agree.

CHAIR WALDROP: Okay, thanks.

MEMBER SANDERS: But it seems reasonable, Insufficient with an Exception.

CHAIR WALDROP: Okay. Comments, questions from the Committee? Okay. All right. Then I think that moves us in the direction of voting on the evidence for measure 1628. Jean-Luc, when you're ready.

MR. TILLY: That's right. The polling is open to vote on evidence. Select 1 for High, 2 for Moderate, 3 for Low, and 4 for Insufficient.

MS. ROILAND: Hi, Amy, if you could text your vote to Jean-Luc now, that would be great. Thanks.

MEMBER SANDERS: I have texted, it's the perils of the new cell phone, I'm all thumbs, apologies.

MS. ROILAND: That's okay. Thank you, though.

MR. TILLY: All right. The voting is one voting High, two Moderate, zero Low, and 21 Insufficient. So the measure is voted Insufficient. So the voting is now open to select if the measure passes with an exception. So select 1 for Insufficient Evidence with exception, select 2 for no exception.

MS. ROILAND: Amy, did you submit your vote via text? And, Bob, did you text Jean-Luc your vote?

MEMBER ARCHULETA: Yes.

MS. ROILAND: Okay.

MR. TILLY: All right. The vote is 24 voting for Insufficient Evidence with an

Exception and zero voting no exception. So the Measure passes Evidence with Exception.

CHAIR WALDROP: Great, thank you. So that moves us on into considering gaps in care, opportunities for improvement and disparities, and also a place where we might look at some of that additional information that Karl mentioned we have now from the Developer. Karl?

MEMBER STEINBERG: Yes. So as far as the gaps, the data that are listed in the original, in the document here, indicate that there is still a significant gap. Like I said, that 2014 study, at least in the VA, indicated perhaps less of a gap with 98 percent mean, but it seems like an important area.

CHAIR WALDROP: Okay.

MEMBER STEINBERG: And it indicates that perhaps having this Measure has resulted in an improvement already.

CHAIR WALDROP: I just want to ask if you might comment, or if Karen wants to, on the Staff's recommendation of Insufficient for the

preliminary rating of opportunities for improvement.

MS. JOHNSON: We went with Insufficient on this just because the data that were provided originally were quite old, they were more than five years. Oh, sorry. We chose Insufficient the first time around because the data that were initially provided were more than five years old, so it didn't really speak to the current gap that might be out there. So, that's why Insufficient. In terms of the papers that were added after the work group call, I might be curious to ask if Karl wants to speak to the gap in the 2014 publication. Is there anything that you want to tell the Committee about that publication in terms of gap?

MEMBER STEINBERG: Well, I mean, this is only for VA facilities nationwide and it was essentially about 6,800 visits, out of which about 6,600 actually had the assessment done, the pain assessment. So, that was -- the mean was 98 percent. So, it seems like maybe at least the VA

is doing it well, but I don't know how much that extrapolates to other outpatient cancer settings and so on. I mean, it doesn't make it any less important.

CHAIR WALDROP: Anything further --

MEMBER SANDERS: This is Amy. I have a comment too when we get to disparities, which is sort of a subset of gaps in care.

CHAIR WALDROP: Now would be good.

MEMBER SANDERS: Okay. So there was a paper in the Journal of General Internal Medicine about a year ago that disparities in -- that collection of data in EMRs for race, ethnicity, and preferred language is terrible. I just have some quick notes here, so I don't have the actual data to provide to you, but, and I realize that reasoning by extrapolation is not exactly rigorous, but if EMRs are not even documenting the basic demographic categories that would define a disparity, then I think it's exceedingly unlikely that pain or anything else is being measured on a regular basis according to some

sort of disparity split group analysis or anything like that.

CHAIR WALDROP: Thank you. Anything else from Karl or Tracy on this? The disparity measure? Okay. Comments from the Committee? Questions? Seeing none, I believe that moves us into the place of voting on gaps and opportunities for improvement. Jean-Luc?

MR. TILLY: That's right. Polling is open, please vote for Performance Gap, 1 High, 2 Moderate, 3 Low, and 4 Insufficient. And I'm sorry, I think that we'll need just one more in the room, if you all could --

MS. ROILAND: Also, Alice, if you could submit your voting, I'd appreciate it. Thank you.

MR. TILLY: And, Bob, if you could send a quick text with your vote here. Thanks. So the results are zero voting High, 15 voting Moderate, two voting Low, and seven voting Insufficient. The Measure passes.

CHAIR WALDROP: Okay. Thank you.

Moving on, we will now consider the scientific acceptability of Measure properties. And we'll start by looking at Reliability and the reliability specifications and then move into reliability testing. Karl?

MEMBER STEINBERG: Sure. So, the Reliability, I mean, this is basically a yea or nay, was the patient screened or not. As you might expect, the kappa is very high, it's 80.81 or, no, sorry, 80.86, so it seems pretty reliable to me. And the people who read it before recommended Moderate and that seems to be appropriate as far as I can tell.

CHAIR WALDROP: Anything from Amy or Tracy? Nothing from Tracy. Amy?

MEMBER SANDERS: I have nothing to add.

CHAIR WALDROP: Okay. Thank you. Let me open it to the Committee. Any comments or thoughts? Bob? Rob?

MEMBER SIDLOW: My one comment on the specification is that, this Measure is sort of behind the times in the sense that it's really

narrowed to stage four cancer whereas in feet on the ground, it's happening much more broadly. So, in terms of its actual usability, I'm wondering about that.

CHAIR WALDROP: Okay.

MEMBER STEINBERG: And we actually discussed that in the work group, but apparently it's a lot more work to bring in other additional, an additional population for this. But we certainly agree in principle.

CHAIR WALDROP: Okay. Other comments? Okay. Seeing none, we can move towards looking to vote on Reliability.

MR. TILLY: So to vote on Reliability for 1628, please select 1 for High, 2 for Moderate, 3 for Low, and 4 for Insufficient. So the results are two voting High, 19 Moderate, two Low, and one Insufficient. The Measure passes Reliability.

CHAIR WALDROP: Okay, thank you. Moving on to consider Validity, we'll look at specifications and validity testing and then any

threats that there are to the validity of this Measure. Karl?

MEMBER STEINBERG: Yes. So, as far as validity testing, initially there was concern that we didn't get the information on how the expert panels conducted their validity assessment and the Measure Developers did produce that information, which is available in the packet. So, I think that's been remedied, its face validity. And, let's see, so that's that. As far as threats to validity, one of the exclusions on this Measure was that people had to survive at least 30 days after a cancer diagnosis.

Personally, I don't know exactly why that would be all that relevant, because you can still screen for pain whether the person is about to die or not, but they didn't say how many of the people died within 30 days, so it might be worth not having that exclusion, but I don't think it's a particularly large threat to validity of the Measure. And I think that's about it. It was considered Insufficient before

by the previous reviewers and I think that the Measure Developers did remedy the lack of information about expert consensus and so on and their methods. So I'm thinking Moderate as far as, that would be my recommendation.

CHAIR WALDROP: Thank you for that. Anything further on the shift from Insufficient to Moderate or anything about? Tracy or Amy?

MEMBER SCHROEPFER: I agree.

CHAIR WALDROP: Okay. Amy?

MEMBER SANDERS: I'm not so sure that face validity, and even by the modified Delphi methods, is really sufficient to bridge the gap between Insufficient and Moderate.

CHAIR WALDROP: Okay. Let me open that for comments. Rob, did you have more to say or was that from before? Okay. Anybody, any other comments?

MEMBER STEINBERG: I just think looking at the algorithm, that -- so if the results are not provided, then it's automatically Insufficient, but since they did provide the

results, that was my reasoning.

MEMBER SCHROEPFER: Well, and I thought that, and correct me if I'm wrong, we said this morning that face validity is considered Moderate by NQF. Is that right, Karen?

MS. JOHNSON: Yes. We do accept face validity. We have some exceptions though, it has to be face validity with a measured score, it has to be systematically assessed, which they've done with their modified Delphi process, and they have to tell us about their experts, how did they do it, and they need to tell us what the actual results are, not just say that it was valid, but give us some details about the results. And I believe they, in their additional materials, they were able to do that.

MEMBER SCHROEPFER: Thank you. Because I do think face validity is a weak validity, but I do feel like they did answer our questions and provide us the additional information we needed.

CHAIR WALDROP: Okay. Any other comments from the Committee? Seeing none, let's

move to voting on the validity of Measure 1628.

MR. TILLY: So this time your voting options have changed a little bit actually because the validity testing is face validity only. Your options are 1 for Moderate, 2 for Low, and 3 for Insufficient, so there is no High option. By all means, if you accidentally -- you can press a different key and it will update your vote, yes.

MS. ROILAND: Alice, could you send me your vote, please?

MS. LIND: I thought I had, I'll try again.

MR. TILLY: So, the voting results are 20 voting Moderate, three voting Low, and one voting Insufficient. So the Measure passes Validity.

CHAIR WALDROP: Thank you. So we'll move on to considering Feasibility or implementation issues of 1628. I'll turn again to Karl, please.

MEMBER STEINBERG: Yes. So, this,

again, it's kind of a yea or nay, was there evidence of a quantitative screening tool, a validated screening tool in the record? It was felt by the reviewers to be moderately feasible. That seems to be reasonable to me.

CHAIR WALDROP: Okay. Anything from Amy or Tracy? Tracy says, no. Amy, any thoughts?

MEMBER SANDERS: No, thanks.

CHAIR WALDROP: Okay. Let me open this to the Committee. Any questions about Feasibility, consideration of this issue? Seeing none, we'll move to voting on the Feasibility of 1628.

MR. TILLY: So, the polling is now open for Feasibility. Select 1 for High, 2 for Moderate, 3 for Low, and 4 for Insufficient. We're actually looking for just one more vote in the room, so if you all could try again. Thank you. All right. And the results are two voting High, 22 voting Moderate, zero voting Low, and zero voting Insufficient. The Measure passes

Feasibility.

CHAIR WALDROP: Okay. That brings us to considering Usability and Use, so the impact, improvement, and unintended consequences of a Measure. I'll go back to Karl, please.

MEMBER STEINBERG: Yes. And since it's a maintenance Measure, I guess we're supposed to have more of a focus on the use and usefulness. As of now, it's not a publically reported Measure, it's not currently being used in an accountability program, although at least the California Department of Healthcare Services is planning to start using this Measure. There were no unexpected findings or potential harms. And it was rated as Low and I don't see anything that suggests, as of now, that would be different.

CHAIR WALDROP: Karen, did you want to comment on the Staff's consideration of this as Low?

MS. JOHNSON: It really does go back to it not being -- can somebody turn off --

MEMBER STEINBERG: Sorry, I'll turn

mine off.

MS. JOHNSON: It really does go back to not being in use and also not having the data to be able to see if there's been improvement over time, which of course is related to not being in use. So that's why we chose Low.

CHAIR WALDROP: Thank you. Any other questions or comments from the Committee?

MEMBER SIDLOW: I have a comment.

CHAIR WALDROP: Rob?

MEMBER SIDLOW: I think that there is a potential negative downstream effect of this, given the 30 day specification and the limitation to stage four. The potential downstream is that an organization might focus too narrowly, which sort of defeats the purpose of, I think, the Measure largely. I support the Low, basically.

CHAIR WALDROP: Okay, thank you. Woody, did you have -- oh, sorry --

MEMBER MOSS: I think that's old.

CHAIR WALDROP: -- it's a card. Any other comments from the Committee about Usability

and Use of 1628? All right. That moves us towards our vote on Usability and Use. Jean-Luc?

MR. TILLY: So, to vote on Usability and Use for 1628, select 1 for High, 2 for Moderate, 3 for Low, and 4 for Insufficient Information. So the results are zero voting High, nine voting Moderate, 15 voting Low, and zero voting Insufficient Information. So the Measure does not pass Usability and Use.

CHAIR WALDROP: Okay. My understanding is that we do continue on an overall vote though, even though it doesn't pass on that measure. So we'll move on to the overall vote of the acceptability of 1628.

MEMBER HANDZO: Yes. I would find it helpful if, potentially, Dr. Lorenz could address this question of usability going forward. That seems to be a critical issue here in terms of NQF. Karen, am I right about that?

MS. JOHNSON: Just a reminder, even though you voted Low for Usability and Use, that is not currently a must pass criterion as

Evidence or Reliability and Validity is. However, that said, our new maintenance process, we are putting more emphasis on Usability and Use. So, again, it's a balance thing. I think hearing from Dr. Lorenz would be actually pretty interesting. Dr. Lorenz, are you still on the line? Or Neil, either one?

MR. WENGER: I am here. I think Karl is better able to address usability with this Measure.

MR. LORENZ: I'm sorry, Neil, if I need to say something, I will need to reattend, I'm holding multiple meetings, but I'm happy to jump in. Just let me know, I'm here now.

MR. WENGER: They're asking about usability of the Measure.

MR. LORENZ: And which Measure are we on now?

MR. WENGER: The same one.

CHAIR WALDROP: 1628.

MR. LORENZ: Oh. Usability of the Measure, I think it has not been used for quality

improvement, if that's the standard.

MS. JOHNSON: I guess maybe our question is, do you have any idea why folks --

MR. LORENZ: Why it's not been used?

MS. JOHNSON: Yes. I mean, folks around the table seem to like the Measure, but it's not being picked up. Any -- just curious --

MR. LORENZ: I think many people just assume that advanced cancer patients are having their pain measured, although it turns out not to be uniformly true. And the vast majority of emphasis now these days in pain management is on reducing opioid use and on minimizing prescriptions in primary care. So, I just think that other than us, there just are not system champions of the cause, per se.

CHAIR WALDROP: Thank you for that. Any other comments or concerns? Christine?

MEMBER RITCHIE: When do we talk about comparability of Measures?

MS. JOHNSON: That would be later.

CHAIR WALDROP: Any other comments or

questions? I think that signifies we're ready to vote on our overall satisfaction with this Measure.

MR. TILLY: That's right, yes. So the question is, does the Measure meet NQF criteria for endorsement? Select 1 for Yes and 2 for No. The results are 24 in support of the Measure voting Yes and zero voting No. So the Measure passes.

CHAIR WALDROP: Thank you. So that brings us to the end of the first candidate Measures grouping. And I'll pass the baton to Sean.

CHAIR MORRISON: So you guys are stuck with me for the next little while. We're going to move to 1617, which is Patients Treated with an Opioid who are Given a Bowel Regimen. Neil, Karl, it's Sean Morrison, I'm not sure which one of you is going to talk about this one as a Developer.

MEMBER CASS: It's actually me. 1617?

MR. LORENZ: Yes, I can --

MS. JOHNSON: Sean, do you mind if I interrupt?

CHAIR MORRISON: I'm sorry, go ahead.

MS. JOHNSON: That's okay. This was in smaller print here, so you probably didn't see it, but actually, Christine, to your point, let's talk about related and competing Measures now for pain. We just went through the pain Measures. There were three of them. And I'm not quite sure -- I have a couple of extra slides. Can you bring those up, Jean-Luc? I'm just going to run through this. This is probably going to be a fairly short conversation, but maybe not. And apologies, Sean, for jumping in.

CHAIR MORRISON: Go right ahead.

MS. JOHNSON: You guys saw this slide this morning, so I just wanted to bring this back to your attention. We've now looked at the two screening Measures from UNC, as well as the one we just talked about, 1628, and we had looked at 0209 earlier today as well, the Comfortable Dying Measure. Let's go to the next slide. So, the

Comfortable Dying Measure right now actually has gone down. So that means that we are actually not going to talk about related and competing with this Measure. If it comes back after our comment period and you guys decide to move forward with recommending endorsement for that Measure after you see the new data that Carol may be able to bring to us, we probably would have this conversation at that point. Okay.

That said, let's go to the next slide. So, pain Measures in the ambulatory setting. Let me orient you to this table really quickly. You have seen and actually have just talked about Measure 1628. It's about screening for pain, the level of analysis is several different things that a facility plans integrated systems, and it's focused on adult patients with advanced cancer. Now, we have some other Measures that are in our portfolio regarding pain in the ambulatory setting.

Now, you guys have not seen these, you know that because you haven't looked at these,

right? So we can't really ask you to talk about what we would call best in class, we can't ask you at this point to say, this one is better than that one or if we had this one, we wouldn't need the other ones, that sort of discussion. However, I did want to point out some of the differences across these Measures. Particularly, I think, in the denominators, and we've gotten to some of this already I believe in the discussion.

What we would like from you guys is, do you see anything on this screen, particularly around the numerator or the denominator or the exclusions, that you could offer as recommendations for what we call harmonization? So, when we talk about harmonization, we talk about making Measures as aligned as they can be, either in the numerator or in the denominator, in the exclusions, et cetera. It is a little tough, most of these are looking at the individual level.

But really I think we just want to have a conversation now, do you see the need for

four ambulatory pain Measures? If you could tweak your magic wand and have something, what would you see? Would you have something different than what you have on the screen in front of you? So, let me just stop there and let's just chat. Christine?

MEMBER RITCHIE: Yes. So, this is what I was alluding to earlier. It seems like the numerators for 1628 and 0384 could be harmonized.

MS. JOHNSON: Okay. If those were harmonized, would we need two of those Measures? Arif?

MEMBER KAMAL: So, I mean, I don't know what the answer is of where we want to go, but 1628 is those with advanced incurable disease on all visits, 0384 is independent of stages, as I read it, but actively receiving therapy. So, meaning the gaps are a stage four patient who is not receiving therapy would not fall under 0384, right? And likewise, a patient not receiving -- a stage four patient -- you know what I mean.

MS. JOHNSON: Yes.

MEMBER KAMAL: So, I mean, I don't know what the true answer is of where we want to be. It seems that we're arguing as palliative care clinicians that stage really has nothing to do with it, so I can see where 0384 seems to make sense, but at the same time, we're not saying that somebody needs to be actively receiving therapy to have a pain assessment either, and that's where I think 0384 falls short. And then 0420 requires the documentation of a plan, which just from a COC, Commission on Cancer accreditation, Joint Commission, and NCCN guideline perspective of what cancer centers are doing, they're not documenting the plan. So that would be a stretch Measure in the current sense for cancer practices.

MS. JOHNSON: Amy?

MEMBER SANDERS: I don't have anything.

MEMBER BERMAN: Sorry, this is the other Amy. Hi. So 0383, I'm wondering if that one might be able to, in terms of the denominator, be the stronger indicator? If we

were to look across these, maybe they could be harmonized toward that denominator because it is more specific about process and multiple processes where that pain would then be assessed and, in this case, have an appropriate care plan. So I'm kind of wondering whether there needs to be a separate screening assessment and care plan, but rather that it is screened at all of those encounters and has the appropriate care plan, the 0383 may be the higher level of saying that we've accomplished those things.

MS. JOHNSON: So, a couple things just to point out, and this is, I think, really to get you guys thinking about having a lot of Measures that are doing pretty much the same thing, right? All of these are looking at pain in an ambulatory setting. They're doing it a little differently in terms of the data source, but I think one of the take home points, and, again, just to get you thinking in this way, by looking at the 0420, you see that it is possible to construct a Measure that looks at assessment and follow-up at the

same time, right? So you don't have to split them out, okay? So that's one thing to think about.

Another thing that you could consider is you can definitely look very closely at, as Amy was saying, the narrow groupings of patients, but 0420 actually doesn't do that, it just says, any patient 18 or older. So, this Measure would work whether you have stage four or not or whether you're on chemo, radiation, or not. So, that's the kind of thing that we want you to think about because what we hear in the field is there's a lot of Measures out there and that it's confusing and can be labor intensive to have to work on different Measures.

So, as overseers of the portfolio, you guys could actually potentially help the field by saying, look, this is what we'd really like to see. Does that mean you'll get what you'd like to see? Not necessarily, different Developers have different reasons. And again, we have some differences in terms of levels of analysis and

data source and that sort of thing, but again, that's the kind of discussion I think that would be helpful to have, because what we want to do is push the field forward. And -- Sean?

CHAIR MORRISON: Yes. I just checked, because I didn't want to trust my brain on this, and I don't know if this is the appropriate time, but 0420 is already in the PQRS program. And so, to Arif's point, it is a Measure that's already in the reporting program that people are being held responsible for reporting. So I'm not so sure given that it really encompasses 1628, 0384, 0383, and takes it a step further, that it doesn't harmonize all three of them across a much wider patient population. And it's already gone through the MAP process to get into PQRS.

MS. JOHNSON: Amy?

MEMBER BERMAN: So, just a clarifying question, does 0420 identify at specific times or could it be one time? I'm looking at the Measure and without any specificity, the kind of specificity for example in the denominator of

0383, I'm just wondering whether one could achieve a low bar of a one-time assessment and have fulfilled the Measure?

MS. JOHNSON: And that is actually a great point, and I don't have the details of the Measure. If you guys were actually looking and evaluating that Measure, we would have that. That's why I'm not asking you actually to do a best in class kind of thing. But that's a good point, and I can look into that and get back to you. I don't know the answer right off. But I think that is a very good point that the screening Measure that you looked at is every visit it should be done.

I'll also point out that even if you guys were looking at all four of these Measures in this project, I actually wouldn't have you try to choose best in class between 1628 and the other Measures because the level of analysis is different and we have, as NQF, said that while we would consider those competing Measures, we would have the same conversation that we always have.

So we just said, we won't even have that conversation. Right now we understand that there's different levels of analysis and different Measures sometimes are needed for that, not always, but sometimes. Does this make sense, the kind of things that I'm asking you to at least think about? Arif, do you have something?

MEMBER KAMAL: No, it does, sorry. And I had to clarify this in my own brain to make it work. So, Patient Reported Outcome Performance Measure can be a misnomer if you focus on the O, and this is actually classified as a Process Measure because it's not actually measuring a change in a patient's health state, right? So I wanted to actually give this more credit than maybe it's due, the 0420 being an Outcome Measure, because we're in this sort of continuous search for an Outcome Measure for palliative care that works for everybody, and I just had to clarify in my head that this is actually very much a Process Measure, none of these are truly Outcome Measures.

MS. JOHNSON: And you know, I did this very, very, very early this morning, so now I'm wondering if it is a PROPM and -- do you remember, Sean? That might be a mistake on my part.

CHAIR MORRISON: I can't.

MS. JOHNSON: Marcia may look it up for us. Nothing is riding on this, so I can get back to you on the details. Just so you know, at some point, as members of our standing Committee overseeing the portfolio, at some point, we may come to you and we may say, here are these Measures and, if you can, pick the best one. And what that means is, you would endorse one and not endorse another, okay? Again, we're not asking you to do that right now, but that is something that we may do in the future. And it's hard, because often folks feel funny about taking endorsement away. Any other questions? I don't want to belabor this part of the discussion, but I think it is valuable to be thinking about the portfolio.

CHAIR MORRISON: Let me just clarify, it is a Process Measure, not a PROPM.

MS. JOHNSON: Apologies for that.

MEMBER CASS: Karen, you said we might be asked to choose one best, but how about combining? Is that another --

MS. JOHNSON: That would be possibly even better --

MEMBER CASS: Better?

MS. JOHNSON: Yes. Because --

MEMBER CASS: Yes.

MS. JOHNSON: -- if you combine, you could potentially reach a broader swath of patients or levels of analysis, et cetera.

MEMBER CASS: Each one has elements that are --

MS. JOHNSON: Yes.

MEMBER CASS: -- better than the other.

MS. JOHNSON: Yes. So, a recommendation for combining in some way would be -- again, it would potentially be a recommendation that might not be taken, but

sending a strong signal from this group, I think is something that Developers would pay attention to. All right. I think we're done with related and competing with pain Measures, at least for now. So now, Sean?

CHAIR MORRISON: All right. So let's go back to 1617, which is Patients Treated with an Opioid who are Given a Bowel Regimen. And again, I'm not sure, Neil or Karl, who would like to describe this one for us?

MEMBER CASS: And actually, it's mine.

CHAIR MORRISON: Actually it's not, sorry.

MEMBER CASS: It's not.

CHAIR MORRISON: It's the Developers. I'm asking --

MEMBER CASS: Oh, I'm sorry.

CHAIR MORRISON: That's all right.

MEMBER CASS: I apologize. I didn't --

CHAIR MORRISON: That's all right.

MEMBER CASS: -- see anybody sitting over there. I'm glad they're here, though.

Thank you.

CHAIR MORRISON: Karl, Neil --

MR. WENGER: Yes.

CHAIR MORRISON: -- who's with us?

MR. WENGER: This is Neil.

CHAIR MORRISON: Hey, there.

MR. WENGER: So this is a longstanding Measure that focuses upon the use of a bowel regimen for patients who are receiving more than one opioid prescription. And I don't actually know what else you want me to say about it.

CHAIR MORRISON: That is a good start, Neil, and I'm sure we'll have questions as we move forward. So, now I have Cleanne and I have Arif. You're on.

MEMBER CASS: Okay. I apologize.

CHAIR MORRISON: We're going to start with Evidence.

MEMBER CASS: All right. Karl's another member of my work group, so it confused me, the two names. So this is 1617, it is a Process Measure, and it's a maintenance Measure.

We're looking at the percentage of vulnerable adults that are treated with an opioid that are also offered or prescribed a bowel regimen or documentation of why it was not needed. So, looking at the evidence, the Developer did provide very good evidence at the beginning when this was first presented.

Evidence from the AGS Panel on Persistent Pain in Older Adults, which is Evidence Grade A, and then evidence from the American Pain Society, which was a strong recommendation. But there is no new evidence and the Developer attests that there's been no changes in the evidence since the Measure was last evaluated. Guidance from the evidence algorithm recommended a Moderate rating on this.

CHAIR MORRISON: And my understanding, Karen, correct me if I'm wrong, unless Arif has major comments, is because there's no change in the evidence, we can actually pass through this and move on to gaps.

MS. JOHNSON: As long as the Committee

agrees that there haven't been --

CHAIR MORRISON: As long as --

MEMBER CASS: Yes, and --

CHAIR MORRISON: -- the Committee agrees.

MEMBER CASS: The Developer offered the rationale that there's a strong relationship between the process of care and positive outcomes for the patient, so I think the evidence is clear. So, moving on to gaps. There was some data presented for performance gap requirements. And the studies were reported to be relatively recently, but they were both more than five years old. There was no data on disparities, but because the data provided was more than five years old, the preliminary rating for opportunity for improvement was listed as Insufficient, even though the data that they did provide suggested that there was significant evidence for a gap and that this could be opportunity for improvement and be used for that.

CHAIR MORRISON: Arif, anything from

you on this one? Open for discussion. Christine?

MEMBER RITCHIE: I'm just looking at the screen, but it looks like the Developers provided updated information?

MEMBER CASS: What's that?

MR. WENGER: There were two studies that had not been completed when we previously submitted this Measure.

CHAIR MORRISON: I'm sorry, Neil, so you're talking about the two studies on, the Hanson study and the Walling's study of 2013?

MR. WENGER: Correct.

CHAIR MORRISON: Great. Yes, so I think that's -- for the Committee, that's on the screen in front of you. The Hanson study found was 250 seriously ill people. The Walling's study was a sample of vets with advanced cancer. And, again, showing 44 percent in the first study and then a rate of 52 percent in outpatients and 71 percent in inpatients, which I think would suggest that there is a gap.

MEMBER CASS: I think our work group felt that there was strong evidence for a gap, that the problem was that the data was considered to be somewhat old.

CHAIR MORRISON: Yes, and I think what you're hearing, Cleanne, is that they've provided updated data from 2012 and 2013, which are only four and three years old.

MEMBER CASS: Okay.

MS. JOHNSON: So, apologies if I confused you guys on how we did this preliminary analysis. The initial submission had several other things that they put in there. The two new things were the Hanson and the Walling's papers. But because the data were from 2007 and 2008, we still considered it to be older data, even though the publication dates are newer.

MEMBER CASS: But the study itself was from 2012?

CHAIR MORRISON: This is the --

MR. WENGER: The study --

CHAIR MORRISON: I'm sorry, Neil. This

is the issue of publication lag, that they're working on data that was older, but the publication lag is 2012, 2013, and I think the Committee has to weigh how recent that actually is.

MEMBER CASS: So, this is included as well in the CMS hospice reporting program and collected through the Hospice Item Set. So, do any members of the Committee know any of the current performance rate in the hospice setting that would be helpful in terms of determining gaps? I think this is a Measure that's widely utilized in the Hospice Item Set. I know in our hospice, we rate about 100 percent on this, because we measure it very carefully, it's easy to use and it's feasible, so it's measured. But I haven't seen the data across the board.

MEMBER TATUM: I guess the other way to say that is, although the data is a little bit old, 2010 is not that bad, I can't say there's a lot of data that the gap's improved by any means.

CHAIR MORRISON: Christine?

MEMBER RITCHIE: Just a brief comment that the denominator for our Measure is cancer patients and not hospice patients.

CHAIR MORRISON: Do I have any other comments on Performance Gap before we move to a vote? Can we move to a vote?

MR. TILLY: To vote on Performance Gap, select 1 for High, 2 for Moderate, 3 for Low, and 4 for Insufficient. Okay. The results are two voting High, 15 voting Moderate, two voting Low, and five voting Insufficient. The Measurepasses Performance Gap.

CHAIR MORRISON: Cleanne, you want to move on?

MEMBER CASS: Yes. Moving on to Reliability, the data source for this is paper medical records. And looking at the specifications, the denominator includes vulnerable adults who are older than 17 years and who have been prescribed an opioid in an outpatient, hospital inpatient, or hospice patient. The work group looked at the

denominator utilizing the term vulnerable patients and the definition given to us is for what vulnerable patients include.

So a vulnerable patient is one that's 75 years of age or older, they score greater than 2 on the Vulnerable Elder Survey 2013, life expectancy less than six months, stage four cancer, or receiving hospice care. The work group had some concerns with the denominator based on, number one, the term vulnerable is sometimes used in other contexts in other settings and could cause confusion, and then that perhaps the denominator excludes certain groups of individuals. And that it would be helpful if the denominator could be palliative care patients or, with the stage four cancer, if it could include all patients with cancer. So we wanted to open that up for discussion with the Committee.

CHAIR MORRISON: I'm sorry, before you go, I keep -- Arif, do you have other?

MEMBER KAMAL: So this is where the

inter-rater reliability seems impressively high for what I think if I were abstracting these charts would be difficult to find. For example, prognostication in arrears, as oncologists, we just don't document that very well at all, so it would be a challenge from an extractor point of view to understand who sort of fits in the denominator. If you presume empirically that you know the answer, then I think the numerator has a high kappa. I mean, I think you can figure out that part. It's just the denominator, which has come up in our calls, was, gosh, in arrears, this is somewhat complex to understand.

MEMBER CASS: Yes.

MEMBER KAMAL: And especially the Vulnerable Elders Survey, I mean, I'm not a geriatrician, I know others are. I don't know how often that's used in clinical practice or an extraction of other things. I know for ACO Measures, the vulnerable elder verbiage is used quite frequently, so it must be a really well accepted respected thing, I just don't know it.

From an extractive perspective, is that an easy thing that you can teach an extractor to find from structured and unstructured data in an EHR? I just don't know, I'm asking.

CHAIR MORRISON: Neil, this sounds like something for you guys to answer over at RAND. Can you help the Committee, both in terms of the denominator and the ease of identifying the denominator and the extraction ease? Because I think this was done in some of your original development work.

MR. WENGER: Correct, Sean. So, the definition of vulnerable, you don't need to have all of the criteria, it's any of those criteria. So, we had very little difficulty identifying stage four cancer patients or patients that have poor prognosis and terminal illness, because there's specific verbiage that was used to identify such cases. The Vulnerable Elder Survey is only used in selected places, though I have to say that it's used in dozens and dozens, maybe hundreds, of different places to identify

vulnerable individuals, but it certainly would not be available everywhere. And then patients receiving hospice care. So, identifying the sample can be done reliably, and that's been shown in numerous studies, but the criticism that it excludes cases that perhaps should be within this Measure is a fair criticism.

MEMBER CASS: So, the work group wondered if the stage four cancer could just be all patients with cancer? And --

CHAIR MORRISON: I hear you, Cleanne, I don't think that --

MEMBER CASS: I know --

CHAIR MORRISON: -- we can rewrite the specs. Much as on many of these we would like to.

MEMBER CASS: Thank you. The other concern we had, just for RAND, was about the exclusions. Excluding non-hospice patients who are already taking an opioid at the time of the measurement period opioid prescription, whether the exclusion was valid or if, again, that could

be a group that could be measured with this, could also be included in this Measure? That's another question for the RAND.

MR. WENGER: I can explain to you why we excluded that group. And it's because we didn't feel confident that we could identify cases where a patient did not need a bowel regimen. When we develop these Measures, we want them to be clinically accurate. And if one was already taking an opioid at the time and they were a non-hospice patient, we thought -- in fact, we were able to prove that outside of the range of the chart extraction, you could be missing exclusions. And we were unwilling to allow for missed exclusions. What this does is limit the sample, but it makes it more likely that you're not incorrectly dinging docs for not using a bowel regimen.

MEMBER CASS: Okay. Thank you. So, the reliability testing is high, because the kappa value, again, related to the numerator, is like 86 percent. And three reliability studies

were presented and there was very good inter-rater reliability provided. So the guidance from the reliability algorithm was that there was a Moderate degree of reliability.

CHAIR MORRISON: Amy, do you have --

MEMBER BERMAN: I have a clarifying question. So --

CHAIR MORRISON: Amy, I'm sorry, could you speak into the mic?

MEMBER BERMAN: Yes.

CHAIR MORRISON: Or use your outside voice?

MEMBER BERMAN: I'm sorry.

(Laughter.)

MEMBER BERMAN: So, I'm just wondering, in the world of oncology, how likely is it and has anybody seen the data on whether or not people are restaged? So, the question is, if they were diagnosed, for example, at stage one, do they actually restage along the way and say that somebody is stage four? And my question goes to this denominator issue, whether or not we

could potentially be missing large swaths of individuals who we intend to measure that potentially would be unknowingly excluded if in fact we either do not restage or don't document the restaging. Question.

MR. WENGER: Is that a question directed at RAND?

CHAIR MORRISON: Why not, Neil, go for it.

MR. WENGER: I would say that that is likely to be a problem, but we did not find it to be a big problem. We do not look specifically for the words stage four cancer, in fact, we almost never look for specific words. We have definitions of stage four cancer and are able to apply those definitions across the board. In fact, we use this Measure here, it's part of our measurement set, and stage four cancer is programmed using natural language processing into our EHR.

CHAIR MORRISON: Amy, you want to follow up? And then I've got Laura.

MEMBER BERMAN: But if I understand correctly, the data that you're relying on is record abstraction, so --

MR. WENGER: It is.

MEMBER BERMAN: -- you are looking at people who have in some way conveyed stage four and my question goes to, do doctors restage? Has anybody looked at the data on that, because I'm hearing about this, but I have not looked at the data, so I'm asking this esteemed body, and if they are not necessarily restaging, are we potentially missing a large swath of people? That's the question. If they are showing up in your record review, I mean, you have your own selection bias in that you're using natural language that identifies people who did. I'm asking about those who don't.

MR. WENGER: Right. No, I'm referring -- the natural language processing is the stuff of today, right? But in all the studies that you're looking at, stage four is identified, but not by requiring the words stage four.

CHAIR MORRISON: Laura, and I may have an answer to that one, Amy, afterwards, why don't we go to Laura first?

MEMBER PORTER: What about using the word metastatic instead of stage four? And I don't know -- I know that this is an issue that's come up for us at the Colon Cancer Alliance about whether or not people, when they're diagnosed with stage two and then they recur, are they restaged? And from what we've seen, they're not. So, they're stage two with a recurrence.

CHAIR MORRISON: So, let me -- Arif, and then I have a question for Neil. Okay. So, Neil, my understanding, because having gone through this with you before, is stage four is in the specs, but in the actual guidance, there are a number of other terms that identify somebody with stage four cancer. Is that not correct?

MR. WENGER: That is correct.

CHAIR MORRISON: So that this is -- in terms of defining the population, it is stage four cancer. In terms of operationalizing this

Measure in real time, there are a number of synonyms that have been used in the description of the Measure that identify people with stage four cancer. Is that a fair summary, Neil?

MR. WENGER: Yes, that's accurate.

MEMBER CASS: So, the reliability algorithm recommended a Moderate rating.

CHAIR MORRISON: So this is -- if there is no change in the evidence since this was initially examined, again, the Committee has the option just to move forward without a vote, given that this has been voted on before by the prior Committee and it was endorsed. And, Arif, I think you still -- do you have a comment? I just couldn't tell.

MEMBER KAMAL: I do. And I applaud RAND for having such great data on inter-rater reliability. I'm just internally struggling with, if we were all presented with a 55 year old with an ejection fraction of 15 to 20 percent, whether we would have consensus within this room of whether their prognosis is less than six

months and thus meets the denominator for the criteria or not. Right. I ask that question. And if we applied CMS hospice eligibility criteria to this patient, then we would need a lot more information than what may be found in -- I mean, you would need sort of a medical director to review that, right, that goes one step above extraction. So, I say, one, I applaud the fact that there is such great reliability data, while at the same time saying, gosh, if I was extracting these charts, I know I would have some trouble.

CHAIR MORRISON: So, I'm back to my question. Is there an objection to moving forward without a vote? Or would people like to vote?

MEMBER SCHWIMMER: I think we should vote.

CHAIR MORRISON: A vote? Okay. So, we're going to vote on reliability of the evidence.

MR. TILLY: All right. So, this is

another one of these particular slides where because only data elements were tested for reliability, there are three options. So, 1 for Moderate, 2 for Low, and 3 for Insufficient. So, again, there's no High option. Okay. And the results are 18 voting Moderate, four voting Low, and two voting Insufficient. So the Measure passes Reliability.

MEMBER CASS: Okay. Moving on to Validity. There was face validity testing only and there were no threats to the validity reported, no risk adjustment, and again, there was a concern about the validity testing having the face validity only and whether the data had all been approved or been provided as to whether or not the mechanisms and the methodology for the face validity testing. But I believe that's all been supplied to us now, is that correct? Okay. So, the Committee initially had said Insufficient because of that lack of information, but the work group felt that if that information was supplied, it would go to a Moderate.

CHAIR MORRISON: Arif, things to add? Amy, are you up for a discussion or are you just a remnant from the last time? She was looking at me --

MEMBER SANDERS: Which Amy?

CHAIR MORRISON: -- like I was crazy. No, I'm sorry, Amy Berman.

(Laughter.)

CHAIR MORRISON: I'll take your comments if you want, Amy.

MEMBER SANDERS: No, I'm good.

CHAIR MORRISON: Okay. Others? Vote, Jean-Luc. Let's do it.

MR. TILLY: Okay. And once again, since there's face validity only, there are just the three options, 1 for Moderate, 2 for Low, and 3 for Insufficient. Okay. And the results are three for Moderate -- I'm sorry, 17 for Moderate, three for Low, and four for Insufficient. So the Measure passes Validity.

MEMBER CASS: So, when it comes to Feasibility, several pieces of information were

provided. One was that while this data in this particular study was paper data extraction, the work group felt that this data could probably be easily extracted from the electronic medical record or the EHR, and wondered what the opinion of the Committee would be with that in that regard.

CHAIR MORRISON: Arif? Committee? Let me ask the Measure Developer. Neil, possibility of having this electronically extracted rather than paper and is this something that you guys are doing?

MR. WENGER: So, I have two answers to that. One is that it is indeed extractable, and in fact, in one of the studies, it was done electronically. The problem is that the exclusions are not yet available electronically until we get the EHR manufacturers to code them in.

MEMBER CASS: Well, I wonder what the Committee -- how much of a burden or a barrier that becomes when there's so much increasing use

of the electronic medical record then? If it's a threat to feasibility?

MEMBER RITCHIE: I think that issue is probably relevant to every single Measure that we're talking about today. And I'm not sure we have a good answer.

MEMBER CASS: Okay. Thank you. So, the guidance from the algorithm was Moderate on the Feasibility.

CHAIR MORRISON: People comfortable going to a vote? Jean-Luc?

MR. TILLY: All right. To vote on Feasibility, select 1 for High, 2 for Moderate, 3 for Low, and 4 for Insufficient. Okay. The results are zero for High, 24 for Moderate, zero for Low, and zero for Insufficient. So the Measure passes Feasibility.

MEMBER CASS: So, on usability, this Measure is already part of the Hospice Item Set and it's been recommended for use in the PQRS and other data sources. So it's already being used in a number of accountability programs. It's not

being publically reported, but the usability seems to be very high and it's already in place.

CHAIR MORRISON: Dr. Kamal? Any questions from the Committee? Jean-Luc?

MR. TILLY: All right. So the polling is now open for Usability and Use. Select 1 for High, 2 for Moderate, 3 for Low, and 4 for Insufficient Information. And the results are three voting High, 20 voting Moderate, and one votimeasure passes Usability and Use.

CHAIR MORRISON: Okay. We go to the overall election.

MR. TILLY: To vote for overall suitability for endorsement, select 1 for Yes and 2 for No. The results are 24 voting Yes and zeromeasure is recommended for endorsement.

CHAIR MORRISON: Well done, folks. Nicely done. So, we're going to go on to the next RAND Measure, which is 1625, Hospitalized Patients Who Die an Expected Death with an ICD that Has Been Deactivated. Again, let me turn things over to Neil first, just for a quick

introduction to this Measure.

MR. WENGER: Measure became important at the time that the ICDs were becoming prevalent. And some early, several reports, mostly case reports and series of cases, demonstrated that they were not being turned off, leading to suboptimal deaths. That was the measure, and we developed it and tested it in only a small number of patients. I have since been told that it's being used elsewhere, though we were unable to find other published data concerning further use.

CHAIR MORRISON: Thanks, Neil. And I have the Cass-Kamal team again. Were you guys doing the same order?

MEMBER CASS: Yes, he's doing it.

CHAIR MORRISON: He's doing it?

MEMBER CASS: He's doing it.

CHAIR MORRISON: All right. Arif, you're up.

MEMBER KAMAL: Okay.

CHAIR MORRISON: All right. Evidence,

sorry.

MEMBER KAMAL: All right. So, yes. So this measure regarding deactivation of an ICD when death is known for patients in the hospital. This is using paper medical records at the facility level. And so, from the evidence perspective, it's mostly highly reliant on one systematic review and several clinical practice guideline statements. And kudos to the NQF staff for updating at the bottom of that page the Heart Rhythm Society and others who have as consensus said that this is an important clinical issue to address. And despite not having much empiric evidence for it, but a lot of clinical support from various groups, cardiologists and palliative care and so on, this was given a Moderate rating for preliminary evidence.

CHAIR MORRISON: Cleanne, anything to add?

MEMBER CASS: No, I'm good with that. I think that there's a good bit of information here that supports the evidence for the

importance of measuring.

CHAIR MORRISON: Open for discussion. Jean-Luc?

MR. TILLY: Okay. To vote for Evidence for 1625, select 1 for High, 2 for Moderate, 3 for Low, and 4 for Insufficient. I think we also need just one more in the room, so if you all could, I'm sorry, but just try and vote again? Okay. The results are two voting High, 22 voting Moderate, zero voting Low, and zero voting Insufficient. So the measure passes evidence.

CHAIR MORRISON: Gaps, Arif?

MEMBER KAMAL: So, this will be a common theme through a couple of the next sections. There's not yet a lot of robust data on ICDs in place that are stopped for hospitalized patients. So the developers submit two studies. The first one of over 700 patients, veterans with advanced cancer, where really there was only an N of one that was sort of eligible for the Measure. And then the second study of

close to 500 patients who died in the hospital, there were 12 who were eligible for the measure and three where the ICD was deactivated. So this was given a rating of insufficient evidence, but overall, I think it actually highlights that we just don't have a lot of data yet with these experiences.

MEMBER CASS: Yes.

CHAIR MORRISON: Cleanne, comments, thoughts?

MEMBER CASS: No, I'm --

CHAIR MORRISON: Christine?

MEMBER CASS: -- I would agree.

CHAIR MORRISON: Discussion from the Committee?

MEMBER HANDZO: Just to -- oh, I'm sorry. Christine, between us, Christine and I had it right, she can't put hers down and I can't put mine up. Just to review, Karen, if we vote insufficient on this, the measure is dead, yes?

MS. JOHNSON: Right.

CHAIR MORRISON: Yes, Debra?

MEMBER WIEGAND: So I'm having angst about this because I do work in cardiology and I know this is a really important issue. And I just presented last week on this topic and people came up to me and said, oh, the ICDs are not being shut off at end of life. So, I'm kind of struggling with, there's a problem here, but we don't have reliable data, so what do we do with that?

MEMBER CASS: One of the things that I've wondered about and was going to hold this comment until further on in this discussion, do we ever recommend that measures go to another panel, like cardiology? There's no crossover here? Because it seems like it's the cardiologists that need to be measuring this.

MS. JOHNSON: People can ask for measures to go to different projects, but we slot things in for different reasons.

MEMBER CASS: I'm sure.

MS. JOHNSON: So, this felt like an end of life --

MEMBER CASS: It's just the off the wall comment, because if they would be having that discussion at the time that they're implanted, it would be much easier for us that are helping them to journey on, having them turned off. It's hard to know when that end point is.

CHAIR MORRISON: Christine?

MEMBER RITCHIE: So, Karen, this does not count as an evidence with exception? No evidence with exception?

MS. JOHNSON: No, that only fits with the evidence sub-criterion. So we're talking about the opportunity for improvement criterion.

MEMBER RITCHIE: Okay.

MS. JOHNSON: Yes.

CHAIR MORRISON: So, I've got Woody and then I've got Paul.

MEMBER MOSS: So, I'm really confused. This is a maintenance measure, which means it somehow survived the light of day three years ago, four years ago, and now we're looking at it

and there's insufficient evidence. So it shouldn't have died in 2012, how is it still alive for us to be reviewing it?

MS. JOHNSON: I think that data were sufficiently recent four or five years ago when they looked at it. Right now, this data is from 2008, 2005, so it's getting old. So, again, the reason that we selected insufficient is just, we don't have more current data in front of us. Five years ago, this data wasn't old, right? Does that make sense?

CHAIR MORRISON: Neil, I know you're not in the room, but I hear an mm-hmm as if you're trying to get a word in.

MEMBER TATUM: I think that was me --

CHAIR MORRISON: Oh, sorry.

MEMBER TATUM: -- and I didn't have --

CHAIR MORRISON: What was that?

MEMBER TATUM: It was me and I didn't have the floor, so I stopped myself. I was next in line, but --

CHAIR MORRISON: Paul, you're next in

line, and then I've got Amy.

MEMBER TATUM: I'd like to make the case that this is important enough and, while it's marked as insufficient, our task is to see if there's a gap. And lack of evidence of a change doesn't mean that this isn't still a gap. And you have one system in which a single person had the defibrillator deactivated and you have another system in which 25 percent had a defibrillator deactivated. So that seems like a gap to me, and I don't know the fact that the evidence isn't updated is so much that we should throw this measure out because I'd make the case there's evidence for gap there.

CHAIR MORRISON: Amy, then Karl.

MEMBER BERMAN: I'm using my louder voice. So, for those who aren't familiar with what this really means, if somebody has an implanted cardio-defibrillator and they are known to be progressing toward death, death is imminent, and they don't have a magnet put on their chest and it isn't shut off, they go

through the most violent of deaths, being shocked repeatedly through the entire death process, oftentimes in front of their family. So, there's a tremendous amount of suffering and a tremendous amount of grief and other kinds of stress as a result.

So, the question, I guess, this kind of goes to the, while it is a rare occurrence, this essentially is, I guess, what you would call a sentinel event. This is the kind of thing that is an absolute. And I'm wondering, even though we have issues with the small data or the time line of the data, but it is in current use, how we might think about this as a group when we know something is so wrong and has been considered the standard and we know that today this still commonly goes on, how we might be able to address this as a group? This is one of those kinds of things.

CHAIR MORRISON: Karl?

MEMBER STEINBERG: Yes. Just real briefly, because I'm saying the same thing

everyone else has just been saying, but even if this happens on a pretty infrequent basis, it's a pretty horrible thing to happen. We have enough evidence just anecdotally that this still happens, to me that is enough to give it a moderate.

CHAIR MORRISON: Gregg?

MEMBER VANDEKIEFT: Clarifying question for Karen, and you may have already said this and I missed it and if so, I apologize, which is, where do we draw the line in terms of the currency of data to call it adequate? I mean, clearly there was a gap and as Paul pointed out, we haven't seen evidence that that gap has been closed. Is there a time line where you say, anything before such and such, we can't use?

MS. JOHNSON: No, unfortunately not. Again, NQF pretty much doesn't give absolutes on many things like that, so, no. I think what you have to think about here is you will have, in some cases, your own personal knowledge of what may or may not be going on, and you can

extrapolate if you want to and say, well, we know this was happening five or six years ago, do I have any evidence that it's changed or that it hasn't changed? And you answer in that way.

CHAIR MORRISON: So, let me just summarize the discussion and then I think we can go to a vote. And I just, I want to have the opportunity for a different opinion. So, what I've heard, and nobody has contradicted it, from Amy Berman, is that many people view the failure to deactivate an ICD prior to an expected death as close to a never event as possible. That the most recent data we have come from 2005, 2006, that demonstrate that 25 percent of people who were eligible to have their ICD deactivated had their ICD deactivated, which means that 75 percent did not. Which is a long way from Amy's one percent.

Before we move to a vote, I just want to make sure that nobody in the room with that level of knowledge strongly disagrees with Amy's statement, because that obviously would make a

big impact on how you view the data, and whether anybody is aware of more recent studies that have a different prevalence rate. And I'm going to ask -- I just wanted to put those two questions out before we move to thinking about this, because what I'm hearing or looking around is sort of a different view from where the NQF staff came down, and I just want to make sure that that was sort of cleared away on the table when people go to vote. Jean-Luc?

MR. TILLY: Okay. To vote on Performance Gap, select 1 for High, 2 for Moderate, 3 for Low, or 4 for Insufficient. And, Amy, I'm sorry, if you could just text me your vote?

MEMBER SANDERS: I did text you my vote.

CHAIR MORRISON: It's in the cloud somewhere, Amy.

MR. TILLY: That's right. It's making its way down here, maybe just one more time just so we're sure.

MEMBER SANDERS: I texted it, but I did it again.

MS. ROILAND: Amy, can you just put it in the chat box? It doesn't seem to be sending through text. Sorry about that.

CHAIR MORRISON: Yes?

MEMBER MOSS: While we're waiting --

CHAIR MORRISON: Sure.

MEMBER MOSS: -- so I just started literature searching and I found a study, Nathan Goldstein, Jean Kutner, from 2010 looking at 900 hospices, 58 percent reported a patient had been shocked with a non-deactivated ICD. I don't know if that's more data or not.

CHAIR MORRISON: It is more data. That's not -- it's a slightly different population, but one you would expect, and I think my name is on that paper, embarrassingly enough.

MR. WENGER: Right. Actually, we were aware of Jean's paper and I guess Sean's paper.

CHAIR MORRISON: It's called Young Onset Dementia.

MR. WENGER: But we didn't think it applied here. But I would love for someone to find a paper that we couldn't.

CHAIR MORRISON: I'm sorry, Neil, you've got to use your outdoor voice, we can't hear you.

MR. WENGER: Oh, sorry. We were aware of that paper, but didn't think that it applied to this population. But would welcome someone finding a paper that we were unaware of.

MR. TILLY: Okay. So, the results are one voting High, 23 voting Moderate, zero voting Low, and zero voting Insufficient. The Measure passes Performance Gap.

CHAIR MORRISON: Arif, you want to move us along?

MEMBER KAMAL: All right. So, under Reliability, so this is abstracted through paper medical records for, under specifications, patients who have been in the hospital for at least three days who then have an expected death. And you can see the criteria there, essentially

receiving comfort care, receiving hospice care, which I presume means inpatient hospice care, had a life threatening disease or expected to die. And the reliability testing, the inter-rater reliability testing, essentially they took ten percent of the 496 sample, the one that had 12 with an ICD in it, and with some bad luck, the 47 that they pulled, none of them had an ICD. So they couldn't calculate an inter-rater reliability test. So that led to the staff ranking it as Insufficient.

And from our work group, there was an open question regarding whether the death had to occur in the hospital. I presume from the abstraction methodology it must, because it says from the facility, but with the ways that hospital related deaths are calculated differently based on who you ask, like UHC, U.S. News and Report, and others record hospital death even if a patient dies in hospice, for example, it's calculated as a hospital related death if it occurs within 30 days. There's multiple

methodologies for calculating hospital related death, how this would fit? And I ask that to Neil or Karl, because we weren't sure exactly how to interpret.

MR. WENGER: So, these were all in-hospital deaths.

MEMBER KAMAL: Okay.

CHAIR MORRISON: Open for discussion.

MEMBER KAMAL: I think as a work group, this is where we had the most thought and cognitive dissonance around what we believed was a good measure, it just doesn't have any sort of quantifiable inter-rater reliability data yet. And, again, it harkens back to the idea that there's not a lot of patients with ICD in these samples, so it may just be reflective of that, or as a work group, we were concerned about whether there was a need to ask for or seek out more data at this time or make that a comment for the next time. And I don't think we had a conclusion about that.

CHAIR MORRISON: Karen, can this be a

comment for the next time? Can you give us some guidance on this?

MS. JOHNSON: It could certainly be a comment. We can try, we have varying success on this, but sometimes we try to highlight certain things when we put reports out for comment, so if somebody out there knows something that we don't know, maybe they will share it. So that might be another option. Just to see.

CHAIR MORRISON: Yes. Taking my Chair's hat off, I know that we will have reliability data on this in about six months. We're just finishing a large multi-site study on AICDs or ICDs and we will have good data within six months. I would be -- if the study falls -- actually, it's closed for enrollment, so it's in the analog -- yes. So, I -- that would be hospitalized patients. I would be comfortable voting for this measure.

MEMBER CASS: It would seem like we need a measure for this one for hospice patients and palliative care, as opposed to hospitalized.

MS. JOHNSON: So, I'm just curious, I'm not a clinician, so I don't know the answer to this, but is it generally pretty easy to find in a medical record if you turned it off? Is this something hard to find or is it pretty clear?

CHAIR MORRISON: It's very, very clear. It's crystal clear.

MEMBER CASS: It's clear.

CHAIR MORRISON: Yes, it's crystal clear.

MEMBER CASS: It's very clear.

CHAIR MORRISON: Yes, I mean, I think if people are concerned, as a clinician working in a hospital, it is very, very easy to find this in the medical record. Jean-Luc?

MR. TILLY: Okay. To vote for reliability, please select 1 for High, 2 for Moderate, 3 for Low, 4 for Insufficient. The results are zero voting High, 22 voting Moderate, zero voting Low, and two voting Insufficient. The Measure passes Reliability.

CHAIR MORRISON: Arif?

MEMBER KAMAL: All right.

CHAIR MORRISON: Continue, sir.

MEMBER KAMAL: So on to validity. So, like would be expected, most of this is around face validity using the modified RAND methodology, using the ASSIST Expert Panel, which these measures come from the ASSIST cohort. So face validity testing was the only validity testing provided. The specific results of the testing are not provided, there's not risk adjustment, and NQF staff provided us with Insufficient, which I think we've seen before, right, face validity only as Insufficient with the group to potentially recommend it to come out of Insufficient, is that right?

MS. JOHNSON: This is one where the RAND group did provide that extra information about face validity, so now they have that --

MEMBER KAMAL: Now we have it? Okay.

MS. JOHNSON: -- information. Yes. So now the question is just, do you have any

concerns about missing data or having meaningful differences? It being a rare event, the meaningful differences may not be as big a deal.

MEMBER KAMAL: Right.

CHAIR MORRISON: Questions from the Committee? Jean-Luc?

MR. TILLY: Okay. And this is a little bit different, so validity testing was just face validity, so there are just three options, 1 Moderate, 2 Low, and 3 Insufficient. Okay. The results are 23 voting Moderate, zero voting Low, and one voting Insufficient. So the measure passes Validity.

CHAIR MORRISON: Arif, on to Feasibility?

MEMBER KAMAL: All right. So this is really about sort of ease of -- well, sort of burden for abstraction. And the developers reported that none of the data elements are included in electronic sources, meaning that while we as a Committee have said that this would be easy to find in terms of nodes, that none of

that data would be structured data that we're aware of, and because of questions regarding burden and paper based manual abstraction, the staff provided us with a Low rating up front.

CHAIR MORRISON: Open for discussion. Paul, I'm sorry, did you -- no, okay.

MEMBER CAUGHEY: Sorry.

CHAIR MORRISON: Yes, Michelle?

MEMBER CAUGHEY: I just have a question. I think these -- the cardiologists are required to have a registry of these and track them very closely, so it just may not be easily abstractable in our literature, but the two data sets could be merged.

CHAIR MORRISON: Christine?

MEMBER RITCHIE: Yes, I'm not sure that it requires that they document that the ICD has been turned off, though. I mean, again, I think this chart abstraction business is sort of what we've been dealing with all along. And I think, to somebody else's point, this is probably easier to find in the chart than a lot of other

things that we're going to be looking for in the chart.

CHAIR MORRISON: Other people, comments, want to address -- Amy, was that you? Yes, it is.

MEMBER BERMAN: Forgive me, but I just want to bring into our conversation, there is a study that was published in 2012, if it is helpful for us, that has maybe a little bit more on prevalence. It's by Tajouri, it is a study of 420 patients, 71 percent were male, who had the implanted cardio-defibrillator. And it talks about a third of those patients having advance directive and only two percent of the advance directives had any mention of deactivating an ICD. In other words, there is a very notable gap in that the cardiologists are not discussing the turning off of their ICD.

CHAIR MORRISON: Thanks. Other comments about feasibility? I mean, I think Christine has raised this issue several times and I think it is real. Are we comfortable going to

a vote on the feasibility? Okay.

MR. TILLY: Okay. To vote on Feasibility for 1625, select 1 for High, 2 for Moderate, 3 for Low, and 4 for Insufficient. And, I'm sorry, Bob, could you try sending me your vote again via text or email? All right. And the results are zero voting High, 21 voting Moderate, three voting Low, and zero voting Insufficient. The measure passes Feasibility.

CHAIR MORRISON: And on to our last, usability.

MEMBER KAMAL: All right. So usability reflects the uptake of this measure and you'll see the note that the NQF particularly looks at uptake into pay-for-performance or accountability programs within three years of NQF endorsement. And this is the story, I think, of multiple stakeholders trying to get this into pay-for-performance systems and it just hasn't gotten there yet. So, you'll see there's planned use for accountability, but not yet. One example is the California Department of Healthcare

Services for a publically reported dashboard, although a final decision has not been made about that.

And then you'll also see the comment regarding the MAP, I'm interested in what Sean has to say about that, recommending inclusion in the PQRS, as well as CAHPS being behind that. And this might fall victim to the fact that I don't think any of our NQF endorsed Measures for palliative care really move forward to PQRS, so this might have just gotten guilty by association. So it was given a Low usability score based on -- I think the major criteria of inclusion in accountability programs within three years of endorsement set the bar.

MS. JOHNSON: That and, again, this catch-22 that we don't have the data, so we can't show improvement over time either --

MEMBER KAMAL: Right.

MS. JOHNSON: -- which is another big piece of Usability and Use.

CHAIR MORRISON: So let me just take

my hat off because Arif asked me to. This was strongly endorsed by the MAP hospital work group. There was a lot of enthusiasm for this measure when it came before MAP and, obviously, I cannot explain the workings and mechanisms at CMS as to what happened after it went from MAP to CMS. But there was enthusiasm and, as said, very strong support, and the MAP echoed many of the comments that we've heard around the room today. I'm so sorry, Linda, go ahead.

MEMBER SCHWIMMER: That's all right. So, this is really a question, maybe, for the clinicians more than anyone else, because I don't know how this would work. If I was a patient that had this implant and I needed to have a conversation with someone to have it deactivated, who would know that I had that, who would initiate that conversation with me, and then, how would that and where would that be done? Because as I've been listening to the conversation, it feels like the measure doesn't necessarily align with the actions that need to be taken and then

would be really difficult to align with either a quality improvement program or a pay-for-performance program. And maybe it's bad luck, but maybe that's why it hasn't really either been studied the way yet we need or been implemented. So maybe somebody could kind of walk me through that. That would be helpful, thank you.

CHAIR MORRISON: Take a crack first, Debra, and then I'll fill in.

MEMBER WIEGAND: Yes, sure. So, I think it would depend on the setting, but in an acute care setting, it's usually part of the advance care planning discussion and then it goes into the advance care planning note that's in the documentation. And then you end up having the whole goals of care, do not resuscitate, they might have some leveled system, but it's usually kind of a red flag and very available to be seen. I'm not sure if that answers the question. And then if somebody is going to go from acute care to hospice, at least from my experience, this discussion happens before they leave, if

everything goes the way it's supposed to, so that when they go to hospice, the device is either deactivated in the hospital or they get them home and then deactivate.

CHAIR MORRISON: And the only thing I would add to that is that the data, at least data from our group, has shown that this is -- even though patients would like to talk, engage in advance care planning, patients with ICDs do not want to talk about this at the time of implantation. And they are very clear about that. And the question is about -- and that has come up, that came up at the MAP as to, is this an appropriate measure at this time or should it have happened a long time ago?

This is one of those areas where patients really do not want to have that conversation at all. And so the question that you raised is, when is the right timing about this? I think that's something that we don't know from the literature, but I would suggest that by the time somebody's heart has stopped

because they're dying, it's probably too late.

MEMBER WIEGAND: Right. And just to add one more thing, on the guidelines that are part of this document, the Heart Rhythm Society and the other cardiology guidelines, do say to have the conversation, at least an initial conversation, pre-implantation, but in actuality in clinical, it doesn't usually happen.

CHAIR MORRISON: As one of our patients said in our focus groups, are you crazy, why would I ever want to turn this off?

MEMBER SANDERS: This is Amy. I know we're not supposed to bring anecdotal data to bear, but in this one instance, I beg that you grant me an exception. In the case of my father, who had an implanted defibrillator, was in the last days of his life, was already comatose, I was a physician, granted I was only a humble intern, but it was the hospice nurse who said, oh my God, we have to turn that thing off. And to this day, 12 years later, I am grateful to that hospice nurse. So I do think that, while as

advanced care planning becomes more common, it may become a common feature of advanced care planning discussions, it is still appropriate at the stage of time that is specified by this measure.

CHAIR MORRISON: Thank you. Very helpful. Jean-Luc, I think we'll go to a vote, unless anybody objects? Okay.

MR. TILLY: All right, great. To vote for Usability and Use, select 1 for High, 2 for Moderate, 3 for Low, 4 for Insufficient Information.

CHAIR MORRISON: Cindi, I'm sorry, did you have a -- that's okay, thank you.

MR. TILLY: The results are four voting High, 19 voting Moderate, one voting Low. So the measure passes Usability and Use.

CHAIR MORRISON: All right, guys. We now go to the final vote.

MR. TILLY: To vote for overall suitability for endorsement to answer the question, does the measure meet NQF criteria for

endorsement, select 1 for Yes and 2 for No. The vote is 24 voting Yes and zero voting No. So the measure is recommended for endorsement.

CHAIR MORRISON: All right, guys. I'm going to actually beg your indulgence. I know we're supposed to take a break at 3:15, but Neil Wenger has been on the phone for almost two hours and we have one more of his measures to do. So, would people be okay if we just did the last RAND Measure and let Neil go back to work? And then we'll take a break at 3: 30? Yes, it's true. We'll let Neil go to lunch. And the good news is, we were over an hour behind and we've now caught up to being just 15 minutes behind. So, well done, folks. Well done. Neil, unless you want to leave, do you want to do your last measure about patients admitted to the ICU who have care preferences documented?

MR. WENGER: Really appreciate it, Sean, thank you. So, this measure is for people who get admitted to an ICU, for vulnerable adults who get admitted to an ICU and are there for at

least 48 hours that there must be documentation about the patient's preferences or documentation why it's not feasible to document about the patient's preferences.

CHAIR MORRISON: Thank you, Neil. And then I've got Amy Sanders and Tracy as the discussants. And I'm not sure if you guys discussed or decided how you were going to go.

MEMBER SANDERS: I think I was designated as the -- this is Amy Sanders. I was designated as the primary respondent and I'm ready to go in that role if that's okay with Tracy.

MEMBER SCHROEPFER: That's okay with me, but can I ask a clarifying question --

CHAIR MORRISON: Absolutely.

MEMBER SCHROEPFER: -- of the developer?

CHAIR MORRISON: Sure.

MEMBER SCHROEPFER: So, Neil, I just want to make sure before we start this, because we were confused when our group had our meeting.

So is this about the discussion? That is, if there's been a discussion about preference in terms of 48 hours? Or let's say that it's me who's in the ICU and I have an advance care plan, would that count even if nobody discussed anything with me?

MR. WENGER: Yes. It simply must exist, the preferences must exist. So if they preceded the admission to the ICU, maybe that's what you're getting at --

MEMBER SCHROEPFER: Yes.

MR. WENGER: -- that would count.

MEMBER SCHROEPFER: That would count? Okay.

MR. WENGER: Yes.

MEMBER SCHROEPFER: Thank you.

MR. WENGER: Yes. I think you'll find for all of our measures that our goal is to make them what would be clinically appropriate under the circumstances, not to check a box.

MEMBER SCHROEPFER: Oh, well, that's helpful, because we really did struggle with that

piece. So, thank you.

CHAIR MORRISON: Amy, we are all sitting in rapt attention. Go ahead.

MEMBER SANDERS: Okay. Thank you. And I thank you also for everybody going ahead with this now, because it will allow me probably to get to the airport on time. So, this is Measure 1626, it is a maintenance measure, it is a process measure. It was originally endorsed in February of 2012. The measure assesses the percentage of vulnerable adults where care preferences are documented within 48 hours of ICU admission. During our working group call, we had a fairly in-depth discussion about definitions of terms. So, as previously, the definition of vulnerable elders, so we can consider that as -- I'll stipulate that as having already been discussed.

There was also some question about what constitutes admission to an ICU. Is it when the patient is accepted by, say, a NICU fellow? Is it when the patient physically rolls into the

ICU? If there could be some comment perhaps added that would clarify that? I suspect that what is meant is that the clock starts ticking when somebody rolls into an ICU, or at least that would make sense to me, but with it not specified, I don't want to assume.

MR. WENGER: So, it's actually when the admission orders get written, and maybe that should be clarified.

MEMBER SANDERS: Okay. So then, in some cases, it may actually be, depending on how long somebody has to wait for a bed, at least in my hospital, I'm not so sure that admission orders and physical presence in the ICU are actually the same thing and they can be separated by as much as hours. In any case, as a maintenance measure, the developer did not present new evidence, but there is, I think, actually some new evidence based on my very quick literature search.

In particular, we're looking to gap in care. The previous evidence included a

systematic review without -- and the processing review. There were some questions about, is it documentation or is it actual review of the preferences? And I think actually I may have misread some of the documentation during the working group call, and if I did, I do apologize for that, because when I read it just now, it felt very different to me. So, thank you, Neil, for the clarification. And I'll stop now and see what I should do next.

CHAIR MORRISON: Thanks. And I'm losing my mind because I can't remember who else I had on this one. Tracy, thank you. Tracy, anything to add?

MEMBER SCHROEPFER: No.

CHAIR MORRISON: So, the floor is open for discussion on Evidence. And NQF's initial recommendation was Moderate. Jean-Luc, vote?

MR. TILLY: So to vote on Evidence for 1626, select 1 for High, 2 for Moderate, 3 for Low, 4 for Insufficient. Okay. The results are zero voting High, 24 voting Moderate, zero voting

Low, and zero voting Insufficient. So the measure passes evidence.

CHAIR MORRISON: Amy, you want to talk to us about gaps?

MEMBER SANDERS: Yes, sure. So, there I believe has already been mention of the Walling 2013 paper that looked at a national sample of VA patients. With respect to this measure, there was an n of 150 and there was an approximately 46 percent gap in care. So, I realize that was just the VA, and I think it -- well, it was just VA patients, but I think there is a gap in care for this measure. And then, my previous comments about disparities hold here as well, that there's evidence in the literature that the very categories that define disparities are not being documented, so I'm pretty sure that this measure is not being assessed according to membership in one or another disparity group.

CHAIR MORRISON: Tracy, anything to add? And, I guess, a question to Karen or Rachel, this was initially by your team rated as

Insufficient. Can you talk to us just a little bit about the reasons for that rating?

MS. JOHNSON: Well, pretty much the same thing. My understanding of the 2012 submission, all the data were five years or older. So it's just a matter of, we don't quite have the data. Is it still a problem, is the question in front of us.

CHAIR MORRISON: Open for discussion. And as Karen said, the reason for staff rating of Insufficient was, is this still a problem? And those who have evidence, rather than anecdote, it would be helpful if we heard it. Debra, I can see you're itching to say something.

MEMBER WIEGAND: I was struggling with my name. I don't have research evidence, but I know from a clinical perspective, it's still an issue. And there's a lot of factors behind that.

CHAIR MORRISON: And, Karen, just to push you guys a little bit, because I know -- one can read a lack of evidence for five years because there are no studies as, A, Insufficient,

or, B, it's still a problem and it continues to be a problem. And I guess my convoluted question is, are we -- is this looking at five years old data without further data, is that a strong problem for us? Or if there is clinical recognition that this is still a problem, do we have to rely on that? Some would say you don't need a randomized controlled trial to see if you throw a brick out the window that it falls.

MS. JOHNSON: Yes. You can rely on your experience, your own knowledge, to answer this question. This is tough, and it's tough for all of these measures, because they're not being used and it's not easy to get the data. It was not easy for us on some of these to put that X on those boxes. But you guys just have to weigh what you're willing -- if you feel like there's still a gap, then you should vote accordingly.

CHAIR MORRISON: If people are comfortable going to a vote -- sorry, Paul, you put your -- fair enough. Jean-Luc?

MR. TILLY: For the actual public

record -- please vote for Performance Gap, select 1 for High, 2 for Moderate, 3 for Low, and 4 for Insufficient.

MEMBER SCHROEPFER: I found a study, but I guess it's too late? Sorry.

MR. TILLY: And the voting is, one voting for High, 22 for Moderate, zero for Low, and one for Insufficient. So the measure passes Performance Gap.

CHAIR MORRISON: Amy, you want to talk to us about Reliability?

MEMBER SANDERS: Sure. And I want to start just by, and I apologize in advance, but I want to revisit the issue of what precisely in terms of documentation satisfies this measure. Because I believe I heard Neil said that the mere existence of advance directives of a POLST or a MOLST or whatever, in the chart, is sufficient, but I direct your attention to Page 3 of the sort of master document, where it says, documentation of having an advance directive, other advance care planning document, or POLST in the medical

record is not sufficient to be counted in the numerator. There it's against the -- this is a distinction to the 48 hour requirement, there's an and for the 48 hour requirement.

But on Page 17, it also says, simply having an advance directive or other advance care planning document in the medical records does not satisfy the criterion. However, a notation in the record during the allotted time period referring to preferences or decisions within such document does satisfy the requirement. So, it seems to me that there has to be some reference within 48 hours of ICU admission to discussion, determination, confirmation, affirmation, reviewing with a proxy, that somehow the decisions about the -- care plans were reviewed with a human being, not just a chart.

CHAIR MORRISON: Neil, can you tackle that one before we go to the group discussion?

MR. WENGER: Yes. I was actually referring to documentation that existed during the admission that was prior to the ICU

admission, not simply having an advance directive in the record. Which these days, the advance directive may very well have been from a prior admission or even in the outpatient setting.

MEMBER SANDERS: Or many years old.

MR. WENGER: Right. But that there needed to be documentation, but that it didn't need to occur during that 48 hour period, is what I was referring to.

MEMBER SANDERS: Okay. I'm not -- how does that square against the statement on Page 3, i.e. the 48-hour time frame also must be satisfied?

MS. JOHNSON: So, Neil, this is Karen at NQF, let me help you out there. That was my wording and perhaps that is confusing. I think what I was trying to get across is that, the documentation, something has to be written, and if I'm understanding it correctly, Neil, something has to be written within that 48 hours, that discussion or whatever has to happen within that 48 hours. It sounds like what you're saying

is that the advance directive or whatever could have been done earlier, but still within the 48 hours, you need to have that discussion. Is that a correct --

MR. WENGER: No. The --

MS. JOHNSON: -- interpretation?

MR. WENGER: No. The way that we implemented this is that there needed to be documentation about preferences that occurred during the hospitalization and you had up to the first 48 hours in the ICU. And in fact, you needed to be in the ICU for 48 hours to even qualify for this measure. Most of the documentation was indeed in the ICU and almost everyone satisfied it that passed it, with documentation in the ICU. But it is true that simply having an advance directive in the record would not satisfy this measure.

MEMBER SANDERS: So, the ICU team needs to speak either to the patient or the patient's proxy to sort of review their wishes upon admission to the ICU. Is that an accurate --

MR. WENGER: I think data in the record beyond that would be acceptable. So, that would certainly satisfy it. But the documentation didn't necessarily need to come from the ICU team, it could be from a continuity doc, it could even be from a social worker or someone from a different team.

MEMBER SANDERS: I'm including everybody who is taking care of the patient while the patient's in the ICU.

CHAIR MORRISON: I'm seeing a lot of puzzled looks around the table. Neil, let me articulate something and tell me if I've got right or wrong. To get into the denominator, you have had to have been admitted into an, be a vulnerable elder by the ACO definition and be admitted into the Intensive Care Unit. For the numerator, therefore, it is the number of patients who, after admission to the hospital on this admission, had their advance directive or proxy designation documented in the medical record or a documentation of their care

preferences and you had up until 48 hours after ICU admission to see that documentation. Is that how I'm interpreting this correctly?

MR. WENGER: That is accurate except it had to be a preference documentation.

CHAIR MORRISON: I'm sorry, preference documentation.

MR. WENGER: Yes.

CHAIR MORRISON: Yes, sorry. And that -- so therefore, you got excluded from the numerator, you were a nay if that preference documentation occurred more than 48 hours after ICU admission or it did not exist at all during that index hospitalization. Is that correct?

MR. WENGER: Correct.

CHAIR MORRISON: Okay. I think -- let me see tent cards, I see a lot of hands.

Christine?

MEMBER RITCHIE: So, just to be clear, the documentation is not necessarily expected during the 48 hours during which they're in the ICU?

CHAIR MORRISON: That is correct. It has to have happened from the time of hospitalization up until the 48 hours after ICU admission. Anything after 48 hours doesn't count.

MEMBER RITCHIE: Right. So, I guess this may -- well, okay, fine.

CHAIR MORRISON: Let's see, I've got Amy and then, is that Tracy there, yes. Amy?

MEMBER BERMAN: So, if someone had an advance directive prior to the hospitalization and if, given the fact that 75 percent of people are unable to make some or all decisions at the end of life, they're admitted to the ICU unable to have their preferences obtained, are you saying that the clinician needs to document what already exists within the clinical record again? Is that -- and if that is the case, I'm just wondering how people feel about an advance directive on file, no ability to reach the family, no ability to have that meaningful conversation, and creating a set of meaningless

documentation. And I'm all for this in concept, but that's what I'm wondering, is that what you mean and is that what we're saying?

MR. WENGER: So, let me make sure that I understand what you're asking. So, there -- let me first point out that most advance directives have little or no usable preference information in them. Then, luckily, that's becoming less and less common, less and less true, but it's still true today. So, there's a surrogate specified and maybe there's a box checked, but little other preference information would one find. But it's true that if an advance directive exists in the record, and there's no reference to that advance directive during the admission or a specification that we have only an advance directive here, no family is available and the patient can't talk to us, which would confer credit for this indicator, then the person would fail, correct.

CHAIR MORRISON: And I just want to remind folks, because we're getting a little

tangential, that we're still on reliability here. So, I'd like to sort of bring the discussion back to reliability.

MEMBER SANDERS: So, if I can follow up that lead, the measure is specified for paper charts. I mean, I think, as we discussed many times, what is abstractable from the paper chart is significantly also abstractable from an EMR. This measure is probably not conducive to reporting merely via structured data elements. And it's also to be reported at the facility level, so we can discuss whether that is the most appropriate level or not.

There was no updated reliability testing performed on this measure. From 2012, there was data from both ACO and ASSIST that showed kappas in the high eighties, 0.87 from the former and I believe 0.86 from the latter. The ASSIST study was limited to cancer patients, this measure is not. And both studies were internal studies from RAND, there wasn't really, to the best of my knowledge, any independent data

presented. And those are my comments on reliability.

CHAIR MORRISON: Open for discussion. Arif?

MEMBER KAMAL: Just real briefly, and file this under not allowing perfection to be the enemy of the good. But alignment with ACO and vulnerable elders makes sense for this specific measure, but as we think about moving forward and making a note, I mean, there's -- a patient with bad COPD with their second time in the ICU would not be in this denominator, right? So, patients with acute respiratory failure on ventilators, unless they have cancer or fall within the VES2, would not. So we'd be missing a lot of really important populations by this denominator. I think we should just note that, I don't think it makes this a bad measure, I think it makes this Measure a little bit more narrow than we eventually want it to be.

CHAIR MORRISON: Other comments on reliability? Yes, Linda? Mic?

MEMBER SCHWIMMER: I mean, just on its face though, there does seem to be an issue with reliability, because I don't know who's deciding whether the POLST form or the advance directive has enough information to be something that could be clinically followed. And unless that's specifically defined, I'm unsure if we really are getting at reliability.

CHAIR MORRISON: Other comments? Thoughts? Jean-Luc?

MR. TILLY: This is again a little bit of a different one, since reliability testing was conducted with just data elements. Your options are 1 for Moderate, 2 for Low, and 3 for Insufficient. So the results are 13 voting Moderate, eight voting Low, and three voting Insufficient. So the measure does not pass reliability.

CHAIR MORRISON: Sorry, guys, we're just checking on the process.

MS. JOHNSON: This would be in the gray zone. So what that means is, consensus not

reached, but we will continue.

CHAIR MORRISON: Amy, Validity?

MEMBER SANDERS: So for Validity, as a maintenance Measure, the question is has any updated information been presented? And the answer there is yes, but no. I do think that Validity needs to be revoted on. The -- sorry. From the previous data, the recommendation for -- there was a, I think, a systematic review and the testing for validity at the data element level was, according to my notes, weak or absent. Hang on, let me look at my other notes so that I make more sense.

So, the ASSIST study included comments from three expert panels. The discussion regarding this Measure was not specific to this Measure. Advance care planning was kind of loosely lumped in with some other Measures related in sort of related areas. At least that was my reading of the study that was added to the shared site earlier this week. And the study was not conducted on ICU patients. So I found that

the data submitted in support of Validity was not as specific as I would have liked for the Measure. The Developers cited originally studies from 2001, 2005, and I think 2011, so much of that data is getting a little antiquated. And I'll stop there.

CHAIR MORRISON: Open for discussion. So, Amy, the NQF Staff had rated this as Insufficient, did your work group have strong feelings to change that in your discussions? Or I'll ask the other members of the group.

MEMBER SANDERS: Not to the best of my recollection.

CHAIR MORRISON: Okay. Jean-Luc?

MR. WENGER: This is Neil. Can I ask a question?

CHAIR MORRISON: Of course.

MR. WENGER: So, the sort of face validity that exists for this Measure is identical to the other Measures that we have discussed before. In other words, structured panel consideration. Actually, two separate

panels, ASSIST and ACO. And then a subsequent clinical panel review of the measures. So, I guess I wouldn't understand why the face validity considerations for this measure would be any different. This Measure also plays a role in the overall process outcome link evaluation that ACO did, because there were substantial numbers of these patients within that cohort. And of course, that cohort showed that process and outcome were linked for both mortality and function.

CHAIR MORRISON: That's helpful, Neil. I think we're just checking some paperwork here because it wasn't clear that the face validity and expert panel review details were provided to us. So I'm just checking with NQF to find out if they have them. And Karen is saying --

MS. JOHNSON: I'm actually going to run it by Neil and make sure we understand, but you provided a document talking about four different -- sorry, the ACO 3, the ASSIST panels with the steps, is that the same panels that looked at

this Measure, Neil?

MR. WENGER: Correct.

MS. JOHNSON: Okay. Then we do have the face validity results.

CHAIR MORRISON: So, I will rephrase that. So we do have the data of the face validity that we applied to all of the other ACO Measures that we've been looking at. So that means that our -- sorry, Amy?

MEMBER SANDERS: I would just -- yes. I would just add though that this Measure is a little bit different. The Measure does not apply to cancer patients only, it's much broader than that. And even though it's, I guess, it's subsumed under the rubric of the ASSIST panel, but to my reading of that document, it wasn't very specifically discussed. So I think that there are potentially some issues regarding face validity with specific reference to this Measure that may not have applied to some of the earlier Measures that we discussed.

CHAIR MORRISON: Fair enough. So I

would say, we should go to a vote and then vote your conscience.

MR. TILLY: Okay. So, again, just three options here since we only did face validity testing. So 1 for Moderate, 2 for Low, and 3 for Insufficient. Okay. So the results for Validity are eight voting Moderate, 11 voting Low, and five voting Insufficient. Which is, I believe, it does not pass.

CHAIR MORRISON: So, at this point, we stop, because it doesn't pass. So, what I would say is, let's take a 15 minute break, we'll come back after five after four. We will work as hard as we can up until our stop time and then, I think we will defer to the morning. Yes? Or are we going to the end, Karen? You tell me.

MS. JOHNSON: We will talk to the folks from ASCO and see what they have to say.

CHAIR MORRISON: And we will -- but let's, everybody looks really tired, this has been a great discussion, so thank you for that, but let's break. Come back promptly at five

after, or be seated at five after.

(Whereupon, the above-entitled matter went off the record at 3:49 p.m. and resumed at 4:05 p.m.)

CHAIR MORRISON: All right, folks. We are back, and we are into the home stretch.

We have modified the agenda just a little bit. We are going to go back and do the two UNC PEACE measures, which will get us probably till about 5 o'clock, maybe a little bit earlier, and then tomorrow morning, we're going to start bright and fresh and early and go through the ASCO measures we didn't get through tonight.

There is -- we will still do -- we will still do just fine. We are cautiously optimistic that we will not need two hours for the CAHPS measures tomorrow, which is what we allocated. And also, I suspect after you guys go home and sleep -- sleep a little bit, you will be -- have sort of gotten into the groove of what this is like, so I think we will be -- we are

continuing to move quickly. We are continuing to gel, but tomorrow morning, we will have really gelled.

So with that being said, let us move on to 1641, Hospice and Palliative Care Treatment Preferences. Dr. Hanson is with us again. Laura?

MS. HANSON: I am back, and I will say that I was very interested in the discussion of the previous measure because I think it illustrates some of the challenges in specifying a measure for this particularly, I think, compelling area of palliative and -- palliative care and hospice.

So the 1641 --

CHAIR MORRISON: I am sorry Laura, I need interrupt just for -- just for everybody else, Christine is recused from this, so she can't play with us in the sandbox for this one. Sorry, Laura, go ahead.

MS. HANSON: The -- I'll miss your wisdom.

The -- the treatment preferences measure is also from the PEACE measure set and was developed for parallel use in palliative care and hospice populations. The measure is designed to capture medical record documentation of communication of treatment preferences for individuals during the initial assessment phase of palliative care for hospice enrollment, and I think I will leave it there unless people have additional questions, and I'll be happy to answer them.

CHAIR MORRISON: Fantastic. So I have as the discussants -- hang on just a second, guys -- yes, I've got Linda and George. I think, Linda, I have you as the lead. Is that right?

MEMBER SCHWIMMER: Right. I am going to --

CHAIR MORRISON: Not George, yes. Linda --

MEMBER SCHWIMMER: Yes.

CHAIR MORRISON: -- go ahead.

MEMBER SCHWIMMER: I'm going to take

the first one and go --

CHAIR MORRISON: Yes, that is perfect. So you want to talk to us about evidence?

MEMBER SCHWIMMER: Yes. So the NQF staff rated this as moderate, and the committee agreed. There was a systematic review, and there was -- although the -- the developers didn't reference -- oh, this is a maintenance and this is a process measure.

Although the developers didn't mention new evidence, there were two pretty recent studies and a lot of publications out there about the importance of -- of documenting your wishes, and there was a lot of -- there was a lot of evidence as well that was supplied by the -- by the developers, so that was rated by moderate, and the committee agreed on that, and happy to answer any questions, or we can keep moving.

CHAIR MORRISON: If there's no objection, given that this is in maintenance, up for maintenance, we can skip forward to gap without a vote.

(No response.)

CHAIR MORRISON: Onwards.

MEMBER SCHWIMMER: Okay. As far as gap goes, this measure on its face looks like it's tapped out, and that's probably because of the -- the definition, so with the -- I am looking right here for my highlighted numerator and denominator, here we go.

So -- so the -- the numerator in hospice situations, in hospice settings, is patients whose medical record includes documentation of life-sustaining preferences, and the denominator is patients enrolled in hospice. So I mean obviously the group that you're looking at already had to have a conversation to be enrolled in hospice, and now you're -- you're looking at how many people have documented their preferences.

So the measure for hospice looks relatively tapped out, but the palliative care, this is one of those ones like we talked about this morning where we're still waiting on

information, and so since that information is coming soon, it is not clear, you know, whether there is room for improvement there or not at this point.

There were some -- some disparities by -- there was some data collected by race, and there were some disparities, so there could be some improvement because there were differences.

CHAIR MORRISON: George, anything to add to that?

(No response.)

CHAIR MORRISON: So again, just because I know people are tired, in the morning we faced this with many of the UNC measures that, on the palliative care data, were not here yet, but they're coming soon, and my recollection was that we accepted that in the morning.

MS. HANSON: And I guess I -- I would just -- I do endorse what Sean just said, and I would just add, although I recognize that it's older data in terms of the gap, the data from the prior submission with the palliative care

population showed that with the serious illness population in the hospital, documentation of preferences 59 percent of the time, that increased to 91 percent after involvement of specialty palliative care, so. And I will just give that to you as part of your consideration.

CHAIR MORRISON: Okay. Questions, comments from the committee? Jean-Luc?

MR. TILLY: Okay, to vote for performance gap on 1641, select high -- or select 1 for high, select 2 for moderate, 3 for low, and 4 for insufficient.

Okay. And the votes are two voting high, 18 voting moderate, and two voting low, and zero voting insufficient, so the measure passes performance gap.

MEMBER SCHWIMMER: Okay --

CHAIR MORRISON: Linda, onwards and upwards?

MEMBER SCHWIMMER: So on reliability, the NQF team gave it a moderate. The only -- the only issue that really -- the only issue that

really came up in the conversation was there was some question about the definition of how you were actually capturing patients' wishes, and -- and we discussed the fact that it's -- it's pretty specifically laid out in the HIS manual, which this is used as part of the hospice quality of scoring, and in that manual, it lays out specifically what is required in a very guide-like sort of way.

And so that's what was -- that's what was used, and there's pretty -- pretty large sample size, and so a good sense of reliability.

CHAIR MORRISON: George?

(No response.)

CHAIR MORRISON: Committee?

(No response.)

CHAIR MORRISON: Jean-Luc.

MR. TILLY: So to vote for reliability, select 1 for high, 2 for moderate, 3 for low, and 4 for insufficient.

The results are two voting high, 20 voting moderate, zero voting low, and zero voting

insufficient, so the measure passes reliability.

MEMBER SCHWIMMER: Again, on validity, this was also graded as -- as moderate, and the committee agreed. Large sample size, there was correlation between the other measures and how this one stacked up, the other measures in the set of six that are used for the hospice quality. There were no questions or concerns on validity.

CHAIR MORRISON: George?

(No response.)

CHAIR MORRISON: Committee?

(No response.)

CHAIR MORRISON: Jean-Luc.

MR. TILLY: So to vote for validity, select 1 for high, 2 for moderate, 3 for low, and 4 for insufficient.

MEMBER SCHWIMMER: So with the exclusions --

CHAIR MORRISON: Hang on just a --

MEMBER SCHWIMMER: Oh --

CHAIR MORRISON: -- second --

MEMBER SCHWIMMER: -- sorry --

CHAIR MORRISON: -- Linda, sorry, I think we're still collecting votes.

MEMBER SCHWIMMER: Oh, sorry.

MR. TILLY: Okay. The votes are zero for high, 22 for moderate, zero for low, zero for insufficient. The measure passes validity.

CHAIR MORRISON: Feasibility.

MEMBER SCHWIMMER: Feasibility. So feasibility, this is currently routinely collected. It is used by CMS as part of their quality measure set, and it's -- according to the developers, it's extracted from the electronic -- the EHR.

CHAIR MORRISON: Team, Karen reminds me that we can carry feasibility and usability forward if we would like, given the other UNC measures that we've agreed to. Any objections?

(No response.)

CHAIR MORRISON: Then we will move those forward.

And I think that brings us to the overall vote, right? All right. Yes or no, red

or blue states.

MR. TILLY: The results are 22 voting yes, zero voting no, so the measure is recommended for endorsement.

CHAIR MORRISON: I am Canadian, so I don't know what red or blue means, so we will move on.

(Laughter.)

CHAIR MORRISON: We're going to move to 1647, which is Beliefs and Values Documentation. Laura?

MS. HANSON: So the last of the PEACE measures that are before you. I will add the caveat that this was -- this was originally a PEACE measure. There was additional work done on it by a different organization that became the measure steward, and then about two months before this meeting, it was passed back to UNC as the measure steward, so I am going to do the very best that I can with this information.

But basically, this is a quality measure to capture the challenge to

hospice/palliative care providers that they address individuals' spiritual needs and concerns during the initial assessment of their needs in various palliative care domains. I will say that one issue that came up in the NQF review that I'll just clarify at this point is that there was a question about whether the measure specifications were modified.

My answer to that question would be no. For the hospice item set, the data-capture period is five days, which is the data capture period that was set by CMS for the initial comprehensive assessment, and so to us, that's not a change in the specifications of the measure, but it is matching the initial assessment period to the period that CMS specified.

CHAIR MORRISON: Terrific, and I think George is taking the lead -- Alice, I must apologize to you on the phone. You're on the second line of my cheat sheet, and I didn't see you as a discussant, so I am assuming that you

will jump forward if you have something to add. Is that all right?

MEMBER LIND: Yes, go ahead.

CHAIR MORRISON: Great. George? Evidence.

MEMBER HANDZO: So this is a direct measure, the percentage of hospice patients over 18 years old who have documentation in the clinical record of a discussion of spiritual/religious concerns or documentation that the patient or caregiver did not want to discuss these concerns, and that takes place within five days of admission to hospice.

This is a process measure here for maintenance review. The -- and oh, let me -- I am sorry, I missed something, back for a second, I failed to say this morning, which I thought I probably should say now, is that part of the evidence for this is the NCP guidelines, and I did serve in the leadership on the -- not NCP, but the National Consensus Conference guidelines, which provided the text, a lot of text for Domain

-- what you now see in Domain 5, so I never worked on this measure, I never worked on NCP, but I do have some knowledge of this.

So there are several guidelines including NCP and NQF performance measures, but there -- it does lack a research literature including any systematic review or QC data, and then I am unaware of any systematic review or expert opinion on this.

The developers do report a new guideline. However, it does not appear that that changes the basic thrust of the guidelines already presented. The measure was originally approved as insufficient evidence with exception. Recommendation of staff is to continue that exception, and that seems reasonable and appropriate for a number of reasons, given that this is the only spiritual care measure currently approved.

There also are, in terms of the exception, there are -- I think one of the characteristics is that there -- it's not a

burden, and there are numerous studies that suggest that patients and family members want to discuss this and that they in fact contrary to -- that not only don't they find it a burden, they find it a welcome discussion to have.

So given that, I would suggest the continuing of the no evidence with -- exception -- insufficient evidence with exception, that's what it is.

CHAIR MORRISON: Alice, Linda, anything to add?

(No response.)

CHAIR MORRISON: From the committee? Arif, is that a card or a finger or is that --

MEMBER KAMAL: It's both.

CHAIR MORRISON: Both.

MEMBER KAMAL: For emphasis.

CHAIR MORRISON: Yes, go ahead.

MEMBER KAMAL: Actually just a real quick question of Laura: so I noticed the NQF -- the press release from 2012 and the Measurement Matters gives this credit of this measure to

DATA, but can you just clarify the relationship between DATA and UNC and the PEACE Project and all that, because --

MS. HANSON: I have no relationship with DATA, except that DATA was the organization that, after the PEACE Project, for a variety of important reasons, they contractually picked up the responsibility to be the measure steward and to do additional testing on this measure, and they maintained that status until about two months prior to this meeting, where they asked to pass this measure-steward role back to UNC.

CHAIR MORRISON: Other comments, thoughts?

(No response.)

CHAIR MORRISON: Guess we'll go to a vote.

MR. TILLY: So we're voting on evidence for 1647. Select 1 for high, 2 for moderate, 3 for low, and 4 for insufficient.

All right. So the results are one voting high, eight voting moderate, and 14 voting

insufficient.

MS. JOHNSON: So my calculator tells me that insufficient is 60.9 percent. That is over our 60 limit, so that actually leads us to the question of does this go forward with the exception?

CHAIR MORRISON: So on the basis of 0.9 percent, we'll go to the next vote.

MS. JOHNSON: Actually, can you hold on just a second? Let me redo that math. What was the -- yes, I think I might have used 24. Let's make sure.

MS. WILSON: I got 61 percent with 23 --

MS. JOHNSON: I was going to say --

PARTICIPANT: -- yes.

MS. JOHNSON: -- it's better than that, so --

CHAIR MORRISON: So with --

MS. JOHNSON: -- yes.

CHAIR MORRISON: -- a whole 0.1 --

MS. JOHNSON: Yes.

CHAIR MORRISON: -- percent more --

MS. JOHNSON: Apologies, I was using the wrong denominator, Christine, I forgot.

CHAIR MORRISON: Let's go to the next vote.

MR. TILLY: Very good. To vote on whether or not an exception will be permitted to evidence, select 1 for -- to pass it as insufficient evidence with exception and 2 for no exception.

So 22 voted insufficient evidence with exception, and one voted no exception, so the measure passes evidence with exception.

CHAIR MORRISON: Onwards to gaps. George?

MEMBER HANDZO: So the developer does provide data on the normally investigated groups for disparity analysis. There are significant differences reported statistically, and I think the question from the staff, I think rightly, is are those -- they're statistically significant. Are they clinically significant?

My -- my sense would be that given the sensitivity of this -- this data, and -- and what we're talking about, I am -- I am loath to dismiss this as clinically insignificant. You know, it just -- I think it deserves a little more work than to just say, well, it's -- it's statistically significant, but we're just going to make a judgment that it's clinically insignificant, and I think we need some more data in order to make that decision.

The -- so the next question is, you know, has this measure topped out? There -- the developers do show that about 20 percent of the population is significantly below the national mean, and we don't really know yet, I don't think, or I don't see the data for what that actually represents.

Are there some disparities hidden in there or whatever? We don't quite know. I mean, I can't figure out from this data, and maybe the developers can add wisdom to that, but I would argue that the -- that there are -- there are

performance gaps here that continue to need to be addressed.

CHAIR MORRISON: Linda, Alice?

(No response.)

CHAIR MORRISON: George, was there a question to Laura in there?

MEMBER HANDZO: No.

CHAIR MORRISON: Okay. Other thoughts, questions from the committee? Arif.

MEMBER KAMAL: So we've done a five-sided study through the PCRC, you know, looking at adherence to this measure specifically, and found that in acute care settings, so ICUs, ERs, it's only 17 percent, and in outpatient and home settings, it only gets as high as 47 percent.

So we have made this a major quality improvement project within our own group, and it's still with a lot of pushing have not seeing closure of that gap, so I think it's an important gap that's still open.

CHAIR MORRISON: I just need to ask Laura a quick question. I mean, my understanding

was the denominator here was just hospice. Is that right? Yes, okay --

MS. HANSON: That is correct. So this -- this is all hospice data, and that in part reflects the transition in the measure steward, but yes, everything presented here is hospice.

MEMBER HANDZO: Yes, I am sorry. The good news for the group here is that we're not having to deal with this hospice/palliative care thing in this measure.

CHAIR MORRISON: Other thoughts?

(No response.)

CHAIR MORRISON: Let's go to a vote. Jean-Luc.

MR. TILLY: To vote on performance gap for 1647, select 1 for high, 2 for moderate, 3 for low, and 4 for insufficient.

The results are zero voting for high, 23 voting for moderate, zero voting for low, and zero voting for insufficient, and the measure passes performance gap.

CHAIR MORRISON: Onwards, George.

MEMBER HANDZO: Reliability: so the measure is specified for the facility level of analysis, and, as we said, hospice in this case only.

The numerator and the denominator are, to my view -- now I will say that I am now into an area that I told people my last statistics course was someplace in the 1980s, so I disclose that this is not my strong suit.

So -- but I think the -- the -- it seems fairly well specified, and the testing there was split-half analysis where the result was 0.94, and a signal-to-noise ratio of 0.99, so the reliability testing seems appropriate. The sample size, as previously discussed, was large, and the staff rated this reliability as high, and I concur.

CHAIR MORRISON: Linda, Alice, anything to add?

(No response.)

CHAIR MORRISON: Anything from the group?

(No response.)

MR. TILLY: Okay. To vote on reliability, select 1 for high, 2 for moderate, 3 for low, and 4 for insufficient.

The results are 12 for high, 11 for moderate, zero for low, and zero for insufficient, so the measure passes reliability.

MEMBER HANDZO: So the validity, this is where the caveat that Dr. Hanson described at the beginning kicks in in terms of the measure being passed back and forth and the staff noting correctly in the -- that this was marked as tested with a modified version, and so there was a question about how it was modified, and does that invalidate the specifications of the measure?

And I think I would agree with Dr. Hanson that it in fact does not. So that was the major question with validity.

Otherwise, the -- the testing is -- is adequate. The kappa score, there's a kappa score which falls into the category of substantial. I

think it's like 0.79 or something like that if I remember correctly. So high enough to be -- to be good.

The threats to validity, the -- the developers tested the persons under 18 years of age and found minimal impact on the result, and so that exclusion doesn't seem to be a threat to validity in this measure.

The measure does appear to be able to identify meaningful differences in performance, and because this is in HIS, the missing data as usual in this measure is minimal. There is no risk adjustment.

So the -- okay, I'm off into validity, so that's the validity -- the validity report.

CHAIR MORRISON: Linda, Alice?

(No response.)

CHAIR MORRISON: Committee?

I don't think you can -- oh, you're not? Never mind. You may weigh in. Please, Christine, speak to us.

MEMBER RITCHIE: I promised you I

wouldn't talk to you about this.

So I guess the question I had to Karen maybe is that this -- this, unless I am misreading it, this received an insufficient -- this received an insufficient, and I was just wondering your thoughts about that.

MS. JOHNSON: Yes. It goes back to that modified question. Basically, what we require is that testing be with the measure as specified, so the testing, the correlation analysis that they used, had that modified on it, so they're testing the modified version, and so we didn't know how different the modified was from this one.

So that was why we had that question. If -- if we're happy that the modified isn't really that different, it's a matter of timing, then we can take the testing as fine.

MEMBER HANDZO: There was just a -- it just appeared kind of in the validity testing, there was this thing, the measure and it said in brackets modified and without specifying what

that was, and so we didn't know.

CHAIR MORRISON: Other thoughts, questions?

(No response.)

CHAIR MORRISON: Should we vote on validity, Jean-Luc?

MEMBER MOSS: So I am confused.

CHAIR MORRISON: Woody?

MEMBER MOSS: You've lost me. So how can we not know how they modified the question? I mean, we're scientists, right? I mean --

MEMBER HANDZO: We do know now. She just -- it wasn't in the report.

MEMBER MOSS: Okay.

MS. HANSON: And I am happy -- I am happy to describe it again.

So basically, the measure specification is that the spiritual -- the assessment of spiritual needs and concerns would happen during the initial assessment by the hospice or palliative care provider. That was how the measure was specified.

CMS then moved in and said there's a five-day window for the initial assessment, which is then basically what's used in the HIS for its implementation. The measure specification itself doesn't change. This is kind of back to the beginning, the morning, in that this is really implementation matching CMS's definition of the initial assessment.

MR. TILLY: Okay, so voting is now open on validity. Select 1 for high, 2 for moderate, 3 for low, and 4 for insufficient.

The results are zero voting high, 21 voting moderate, zero voting low, and two voting insufficient. The measure passes validity.

CHAIR MORRISON: And George, wrap it up with feasibility and usability.

MEMBER HANDZO: Feasibility: all the data elements are routinely generated and available in electronic form. While this measure would seem to be -- I'm going to talk a little bit -- I will say something about that in -- about obvious care delivery potential, the

developers don't report on that, nor are they really required to.

So it is -- it would seem essential, and there is a rationale for having a continuous measure which I think is legitimate.

CHAIR MORRISON: And I am told by Karen that again, we can carry this over from this morning if there are no objections. It doesn't need a vote.

MS. JOHNSON: And let's make sure that I'm not mistaken. This is one of the HIS measures. It's collected in the same way --

MEMBER HANDZO: Yes.

MS. JOHNSON: -- used in the same way --

MEMBER HANDZO: Yes, that is accurate.

MS. JOHNSON: Okay.

CHAIR MORRISON: So we can move on to usability. George?

MEMBER HANDZO: Yes. There do not seem to be any related or competing measures here. The -- the developers do report some data

which suggests that more patients who had a documented discussion had improved their spiritual distress scores. They do report that they cannot yet have sufficient longitudinal data to report on improvement of use. They also report no unexpected findings and no unintended consequences.

I would simply add here that this is the only spiritual care measure, and it is I think critical to what Dr. Hanson started out with about this being kind of in a -- in a field by itself, that it is out of this data that we're going to be able to start to generate data about who wants to have the discussion, what do they discuss, what are the issues they raise, and that's going to drive our interventions, so it's off this data that we're going to start to be able to get a handle finally on this whole area of care which we do not have presently evidence for.

CHAIR MORRISON: Alice, Linda, anything to add?

(No response.)

CHAIR MORRISON: Committee?

(No response.)

CHAIR MORRISON: And again, we can carry forward this morning for the -- unless I hear any objections.

Hearing none, we will carry that forward, and we'll go to the overall vote on endorsement.

MR. TILLY: So to answer the question does the measure meet NQF criteria for endorsement, select 1 to vote yes and select 2 to vote no.

The results are 22 voting yes, one voting no. The measure is recommended for endorsement.

CHAIR MORRISON: Thank you, guys. That was perhaps the smoothest and least contentious discussion on spirituality in palliative care I have ever experienced at any time.

So with that, I think we are done with

the measures for today, so I'm going to turn it over to Karen. We're going to talk -- do the related and competing discussion, and then open it up at the end for public comment, and then we'll talk a little bit about tomorrow.

MS. JOHNSON: Okay. Jean-Luc, if you will just pull the very last slide in our deck. I think this is actually going to go very quickly, but then again, I was told that I jinxed us this morning by saying that we would get -- so I won't say that, I am sorry, just forget I said that.

These are the measures that I have pulled together for us to talk about, not as competing measures, but just related measures that are doing similar things, and I wanted to get your feedback on these.

However, Measure 1626, the care preferences -- no, I'm not sure, which one did we just do? We just did 1641. So 1626 is the one that we did not put forward as suitable for endorsement.

So just pretend that first column isn't there. Because we did not find it suitable for endorsement, we will not be having this discussion about related and competing.

So again, I am not asking you to make any decisions. I really wanted to just let you see a little bit of the difference between the treatment preferences measures that we talked about a couple minutes ago and another measure that we have in the portfolio. It's advance care plan.

The advance care plan measure is -- it actually hits pretty much all settings, and I didn't put all the lists there. There might be one or two that it's not there, but it's really specified for a lot of different places. The same level of analysis, clinician/individual/group practice.

The patient population is much broader. It is looking at all patients 65 and older. But the numerator is -- is quite different, so it's a documented advance care plan

or surrogate or documentation of discussion.

So I think my question to you is as overseer of the portfolio, how -- can you explain how the -- the two measures are different, the treatment preferences versus advance care plan, and would you have any suggestions about, you know, is this something that one could combine and make one measure that does everything? And possibly not. So I am just wanting your feedback on -- on this.

CHAIR MORRISON: And the cards go up. Okay.

(Laughter.)

CHAIR MORRISON: I've got Michelle, I've got Woody, I've got Karl, I've got Amy. Michelle -- and Paul. Michelle, start us off.

MEMBER CAUGHEY: Thank you.

We are in the process of trying to implement both of these things for four million people, and they are very different, so that we would encourage everyone to have an advance care plan and to designate a surrogate, and that's

very different but related to patients who then subsequently or at certain times of life or, if possible, related to their age, but not quite as necessary, and you have a setting definitely in the hospice and hospital where you have then documentation of life-sustaining treatment.

So they are related, but I think we need to see that this process is a stepped process as we move patients along in their journey of all of us through life, and that the very first thing is to designate the surrogate, and then when we're comfortable and we -- and we should move on to that second step or third step.

CHAIR MORRISON: Woody, you were almost as aggressive as Michelle with your card.

MEMBER MOSS: So -- and unfortunately, I disagree with Michelle. I think this is a stepped process, but it's a spectrum, it's a continuum, and in fact, I've been giving talks ever since Medicare and insurers started paying for advance care planning.

We start with the 18- or 19-year-old

who might want to designate who they would want to have as their decision-maker, but they also may have values that they would want to do a living will as well as a durable power of attorney for healthcare, and then as they develop a life-limiting illness, they may want to revisit things, but then if you wouldn't be surprised if they died in the next year, which is sort of further along in the spectrum, then you might want to then more specifically talk about what's in a POLST form and their life-sustaining treatment preferences.

But remember, if they've done a living will at the age of 21 saying I wouldn't want to be -- you know, have my dying prolonged if I'm dying or permanently unconscious, they're already stating some preferences with regard to life-sustaining treatment.

So I think these two can be lumped if we have a broader understanding of what advance care planning is, and I really would like to see us get rid of most of the restrictions. I mean,

it doesn't have to be -- you don't have to be over 65. You don't have to be in hospice or receiving specialty palliative care.

I think we believe it's part of good preventive healthcare to start having discussions with people. It's patient-centered care. It's just finding out what are their preferences, what are their life experiences, what did they see their mother go through, their father go through, their grandmother go through?

We start documenting, and as Michelle suggested, we come back and revisit as they proceed along their trajectory of their life, but I would love to see it as one broad, comprehensive thing that can be done on a repetitive basis over time.

CHAIR MORRISON: Karl?

MEMBER STEINBERG: So I partially agree with Woody. I think, you know, that last measure, why isn't it all patients over 18? That's what's really recommended. Everybody should, you know, document basic preferences and

designate a person. That's recommended for everyone.

But I also agree with Michelle. I think there are two very different things, and I -- I mean, everybody doesn't need to be deciding do they want a feeding tube, do they want CPR, do they want the ventilators, do they want, you know, transfusions or dialysis and that sort of thing. That's really for a select population.

So to me, I don't think these two could be -- could be lumped.

CHAIR MORRISON: Amy.

MEMBER BERMAN: I believe that these are two very different things as well, and while I do believe that they are on a continuum, I think that it is akin to the difference between palliative care and hospice. You can have palliative care within hospice, but palliative care is bigger than hospice.

And here, what we're talking about in terms of treatment preferences, this happens on diagnosis and is not necessarily related to end-

of-life care. This is a broader spectrum of, you know, do we understand what's important to you? And the buckets are independence, function, pain, length of life, and they're going to give it to you in their real words, but that is what the data would suggest. You know, those are kind of the buckets of treatment preference.

Advance care planning really is much more specific about end-of-life care and leads to advance directives, so these really are two different things. I will give you an example from myself. I have treatment preferences that have nothing to do with my end-of-life care that have led to my having very good care, and palliative care is a backbone of that. But my advance directives, I have a healthcare proxy, and my family is well aware of what my wishes are at the end of life, as are my clinical team. So very different things.

CHAIR MORRISON: Paul, are you up or are you down?

MEMBER TATUM: No, I have very little

because everything I wanted to say has been said, with the exception I agree with the differences between advance care planning and treatment preferences. I think as one gets closer to the hospice/hospital/end of life, you know, upstream treatment preferences shift, but I wanted to speak strongly in favor of 326. One of the things we've been missing is some of the primary care component, and the real strength of that is role of the primary care on that measure.

And the -- you know, the danger of relying too much on the 1641 is the importance of the primary care upstream, and the 326 takes it a little more strongly, but I think there is a difference, and I think as you get close to that shifting preferences that can happen, it is discrete where the 1641 has value.

CHAIR MORRISON: Deborah.

CHAIR WALDROP: Okay, so my two thoughts are -- my two cents and my two thoughts are first of all that people die in years, they die in months, they die in weeks, they die in

hours, they die in days, they die in minutes, and when the focal length changes so do people's wishes often, and I think it is really important -- I feel strongly that these are very -- two very separate processes, as Michelle has said, and that to blur them would take away how the focal length really shifts people's preferences.

And the second thought that I want to share is work that Andrew Billings and Rachelle Bernacki recently published on the Goldilocks phenomenon, that you can have advance care planning that is done well too far in advance, like, you know, it's too hot; they can -- you can have it too cold when it's very close to the very end; or you can have it just right, like Goldilocks.

And for me, it's the just right, the tipping point, that we're looking for. And so I think it's very sensitive, and we need to be able to see them as very separate processes.

CHAIR MORRISON: Cindi?

MEMBER PURSLEY: I just wanted to

clarify that for treatment preferences from a hospice perspective, do you want to have antibiotics? Do you want hydration? Those are the types of things. It's not -- they're very separate from advance directives and CPR and proxies and those kinds of things. These are truly about the -- the treatments for comfort in the last days, weeks, or months of life. And so it's a -- to me, they are completely and totally separate.

CHAIR MORRISON: Karl, you get the last word.

MEMBER STEINBERG: Woo-hoo.

CHAIR MORRISON: That was a good last word.

So Karen, I think there was some strong opinion. I hope you got it. Anybody else want to weigh in on this one?

(No response.)

CHAIR MORRISON: So if not, then we now turn to public comments, and first from the phone lines, so if I could ask the operator to

open up the phone lines for public comment?

THE OPERATOR: Okay. At this time, if you would like to make a comment, please press star, then the number 1.

There are no public comments at this time.

CHAIR MORRISON: Thank you. And then I look to those folks in the room, and if you could come up to the mic, that would be terrific.

MS. SINCLAIR: Hello, my name is Stacie Sinclair, and on behalf of the Center to Advance Palliative Care and our partner organizations within the National Coalition for Hospice and Palliative Care, we wanted to thank the standing committee for the excellent work that they have done thus far.

Many of the points that you deliberated are well taken and speak to the gaps that we have in the field. We do want to reiterate, however, that there are very few NQF-endorsed palliative care and hospice measures. This not only has implications for quality

improvement but for the thousands of palliative care providers in whether they can be adequately measured under the burgeoning pay-for-performance programs.

In order to move the field forward in quality measurement and get to a place where next-level measures can be developed and considered, it is critical to maintain the foundation that our measure developers have spent many years building. Therefore, we just wanted to voice our support for protecting the endorsement of all the quality measures that we currently have, including 1639, 0209, and 1626. Thank you very much.

CHAIR MORRISON: Thank you, Stacie. Others?

(No response.)

CHAIR MORRISON: So just before I turn things back to Rachel, I wanted to express both Deborah's and my thanks for a very productive and, I know, long day. I will tell you my experience, this gets easier, and this is a

committee that is really wrestling with some very difficult measures in an area where there are very few data, but it is critically important to get it right, so we all recognize how hard this has been today and thank you for that.

As you notice looking at the agenda, we are a little bit behind, but that is okay. NQF always seems to have a very aggressive plan for day one. However, in the interest of people getting to flights and keeping the anxiety in the room down just a little bit, since we don't have any measures on the docket to talk about anxiety, what I'd like to propose is that we keep breakfast available at 8, but if we could start 15 minutes earlier and have people ready to go at 8:15, we'll try to keep the introductory remarks short, and then we can move pretty quickly forward on the ASCO measures and then on the CAHPS measures and have everybody out on time, if not sooner.

So thank you all very very much, and Rachel, you have some other things to add,

correct?

MS. ROILAND: Just a few things. So I just want to echo Sean's thanks. I think all of us at NQF just want to say thank you for your sticking with us all day today. And just, like Sean said, we'll be starting at about 8:15 tomorrow morning.

And then for tonight, though, we do have dinner reservations for the group at P.J. Clarke's, which is at 1600 K Street, and it's just about a block, block-and-a-half from here, so you probably do have time to go back to your hotel and walk there. The reservation is at 5:45, so if you want to do that, you're more than welcome to just go there individually, or I'll walk a group over, if you just want to hang around here, I will walk you over around 5:45 as well, so it's -- we have about an -- well, little under an hour, so --

Did you have any questions?

MEMBER HANDZO: Tell me the address again.

MS. ROILAND: Sure. It is 1600 K Street, NW. And just tell them you're there for the reservation for Roiland and the National Quality Forum.

Well, we're actually down in the Sidecar. It's a special room downstairs, so we tried to treat you well and get you a special room, so P.J. Clarke's.

CHAIR MORRISON: So thank you everybody, and we will see everybody bright and early tomorrow morning.

(Whereupon, the above-entitled matter went off the record at 4:55 p.m.)

C

M

O

Q

T

W

C E R T I F I C A T E

This is to certify that the foregoing transcript

In the matter of: Palliative and End-of-life Care Steering Committee Meeting

Before: National Quality Forum

Date: 05-10-16

Place: Washington, DC

was duly recorded and accurately transcribed under my direction; further, that said transcript is a true and accurate record of the proceedings.

Neal R Gross

Court Reporter