**NATIONAL QUALITY FORUM**

**+ + + + +**

**ALL-CAUSE ADMISSIONS AND READMISSIONS STANDING COMMITTEE**

**+ + + + +**

**THURSDAY**
**JUNE 9, 2016**

**+ + + + +**

**The Standing Committee met at the National Quality Forum, 9th Floor Conference Room, 1030 15th Street, N.W., Washington, D.C., at 9:00 a.m., John Bulger and Cristie Travis, Co-Chairs, presiding.**

**PRESENT:**
**JOHN BULGER, DO, MBA, Chief Quality Officer, Geisinger Health System, Co-Chair**
**CRISTIE UPSHAW TRAVIS, MSHHA, Chief Executive Officer, Memphis Business Group on Health, Co-Chair**
**KATHERINE AUGER, MD, MSc, Assistant Professor of Pediatrics, Cincinnati Children's Hospital Medical Center**
**FRANK BRIGGS, PharmD, MPH, Vice President, Quality and Patient Safety, West Virginia University Healthcare ***
**JO ANN BROOKS, PhD, RN, Vice President of Safety and Quality, Indiana University Health System**
**MAE CENTENO, DNP, RN, CCRN, CCNS, ACNS-BC, Director, Chronic Disease Care, Baylor Health Care System**
**HELEN CHEN, MD, Chief Medical Officer, Hebrew SeniorLife**
**WILLIAM WESLEY FIELDS, MD, FACEP, Assistant Clinical Professor, UC Irvine Medical Center; Board of Directors, CEP America**

**PAULA MINTON-FOLTZ, RN, MSN, Assistant Administrator, Education, Patient Safety and Quality, Harborview Medical Center - University of Washington**
**BRIAN FOY, MHA, Vice President, Product Development, Q-Centrix, LLC**
**LAURENT GLANCE, MD, Vice-Chair for Research, University of Rochester School of Medicine**
**ANTHONY GRIGONIS, PhD, Vice President, Quality Improvement, Select Medical**
**BRUCE HALL, MD, PhD, MBA, Professor, Surgeon, Washington University; Vice President for Patient Outcomes, BJC Healthcare**
**LESLIE KELLY HALL, Senior Vice President, Policy, Healthwise**
**PAUL HEIDENREICH, MD, MS, FACC, FAHA, Professor and Vice Chair for Clinical, Quality, and Analytics, Stanford University School of Medicine, and VA Palo Alto Health Care System**
**KAREN E. JOYNT, MD, MPH, Assistant Professor, Brigham and Women's Hospital**
**KEITH LIND, JD, MS, BSN, Senior Policy Advisor, AARP Public Policy Institute**
**CAROL RAPHAEL, Senior Advisor, Manatt Health Solutions**
**PAMELA ROBERTS, PhD, MSHA, ORT/L, SCFES, FAOTA, CPHQ, Manager for Inpatient Rehabilitation; Quality, Education, and Research; and Neuropsychology, Cedars-Sinai Medical Center**
**DEREK ROBINSON, MD, MBA, FACEP, CHCQM, Vice President for Quality and Accreditation, Health Care Service Corporation**
**THOMAS SMITH, MD, FAPA, Medical Director, Division of Managed Care, New York State Office of Mental Health; Special Lecturer, Columbia University Medical Center ***

**NQF STAFF:**

**HELEN BURSTIN, MD, MPH, Chief Scientific Officer**
**ELISA MUNTHALI, MPH, Vice President, Quality Measurement**
**TAROON AMIN, MPH, Consultant**
**SHEILA CRAWFORD, Administrative Manager**
**DONNA HERRING, MPH, Project Analyst**
**ERIN O'ROURKE, Senior Director**
**ZEHRA SHAHAB, Project Manager**

**ALSO PRESENT:**

**HAIKUN BAO, PhD, Yale-CORE ***
**KYLE CAMPBELL, PharmD, MS, Health Services Advisory Group, Inc.**
**ELIZABETH DRYE, MD, MS, Yale-CORE**
**TERRY FISHER, MPH, PMP, MD Anderson**
**BARBARA JAGELS, RN, MHA, CPHQ, Seattle Cancer Care Alliance**
**NICOLE KEANE, MSN, RN, Abt Associates**
**ALAN LEVITT, MD, CMS ***
**KASIA LIPSKA, MD, MHS, Yale-CORE ***
**STEPHEN MCKEAN, PhD, CMS**
**DENISE MORSE, MBA, City of Hope ***
**KAREN PACE, PhD, RN, Health Services Advisory Group, Inc. ***
**JENNIFER RIGGS, MS, PhD, Abt Associates**
**ERICA S. SPATZ, MD, MHS, Yale-CORE**
**SUSAN WHITE, PhD, RHIA, CHDA, The James Cancer Hospital-Ohio State University Wexner Medical Center ***
**ALMUT WINTERSTEIN, RPh, PhD, FISPE, University of Florida**

*** present by teleconference**

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

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

9:08 a.m.

CO-CHAIR BULGER: Okay. We're going to get started here for day two. Welcome back. Thank you again to Bruce for working us through this stuff yesterday while I traveled down here. I missed Dr. Krumholz yesterday morning, but I got to see his tweets last night once I got on the Twitter, so I know everything he said. Although he missed a "not" in one of his tweets and he said he -- there was some comment about thinking that there should be a double standard instead of I do not think there should be a double standard for SES.

(Laughter.)

CO-CHAIR BULGER: I was thinking about correcting it, but I let it alone.

CO-CHAIR TRAVIS: Oh, then he fixed it.

CO-CHAIR BULGER: Oh, he fixed it.

CO-CHAIR TRAVIS: I emailed him, yes.

CO-CHAIR BULGER: All right.

CO-CHAIR TRAVIS: I didn't think he wanted to --

(Simultaneous speaking.)

CO-CHAIR BULGER: Yes, I saw that and I was like I don't think that's what he wanted to say.

Welcome also to the public and anyone else who's on the phone. This is the second day for the Readmissions Project Standing Committee for those that aren't sure where they are.

So we have a host of measures this morning. The first three this morning are very similar, so I think we'll probably end up having the same type of discussion we had with some of the other ones yesterday, where the first one there's a lot of discussion and then the following two are very similar to that.

So we'll get right into it. The first measure is 2886. It's a new candidate measure, Risk-Standardized Acute Admission Rates for Patients with Heart Failure. And as I said, there's this one, which is heart failure. The

next one is diabetes and the one after that is AMI.

So we'll start with the developer.

DR. DRYE: Hi, I'm Elizabeth Drye and one of the directors at the Center for Outcomes Research and Evaluation at Yale. I'm also a pediatrician and I've been working on developing outcome measures for a long time with our wonderful team.

I was going to provide some background remarks on these ACO admission measures just to give you an overall picture of their differences and similarities with the measures you looked at that CORE developed yesterday. And the Erica Spatz who's a cardiologist and led the development of the first measure that is on the docket, the heart failure measure, will give you a couple more detailed remarks about that measure, if that sound okay. I'll try to be quick.

So thanks again for the opportunity to introduce these measures. We're really excited

about these measures because they move from looking at hospital care for patients who are acutely ill to looking at outcomes for the management of ambulatory care patients with these chronic conditions: heart failure, diabetes. And the third measure is for patients with multiple chronic conditions. They're similar in their approach to the measures you looked at yesterday in that they share some of the same modeling and risk-adjustment strategies. And also we test validity and reliability and disparities in the same way, but I'm going to focus now on their differences.

The biggest difference is that we're measuring ambulatory care quality over the course of an entire year for non-hospitalized Medicare fee-for-service patients. And the measures are really looking -- the measure concept is evaluate how well the providers are working with each other and with their patients to improve outcomes.

The outcome we use is acute unplanned

admissions per 100 person-years, which is a count of admissions. And the reason that we're using acute unplanned admissions is we are evaluating how well providers are working together to avoid catastrophic complications of the chronic conditions that are the focus of the measure, but also lower the overall risk of a hospitalization for these generally more vulnerable patients through providing best preventive care, early intervention in acute exacerbations of illnesses and also avoiding complications of chronic disease management such as adverse drug events, drug interactions, etcetera.

Second, we're in a very different program context, which is exciting for us to be pioneering. These are accountable care organizations. And actually these measures are already in use in the ACO, CMS' ACO quality measure set. They'll be reported for the first time this summer. They didn't way for NQF endorsement, but they're very interested obviously in getting endorsement. And they are

reported over a calendar year and we use data from the prior years to develop our patient cohort and evaluate risk-adjusters. And they'll be reported on 2015 performance within the next month or so.

The Medicare Shared Savings Program and the other ACO Programs at CMS Pioneer, and that's evolving into the next gen programs, use these measures. These are very different than inpatient quality reporting in that they're voluntary programs. Providers come together and opt to participate and take shared responsibility for the Triple Aim essentially of providing better care, better health. So population health management and lowering costs. And they're jointly responsible for the outcomes of care, so that makes it a wonderful environment to be evaluating population-based outcomes.

As a result of the way the programs are structured and their goals, our conceptual approach to these measures is it takes an expansive view of how providers might act to

lower the risk of acute admissions. We have a conceptual model that for this first measure is on page 42 of the testing part of the form, but we lay out there; and this is relevant to how we think about disparities and risk-adjustment, a really broad range of factors that I think everyone is aware of that could influence the risk of admission including health factors, health behaviors in the environment, environmental factors and community resources, etcetera.

And we had a lot of discussion with this about our expert panel and with CMS and we took public comment on it and decided that we would focus our risk-adjustment on patient health status and age at the outset of the near and not adjust for these factors, because broadly within the ACO community there is much innovation going on to mitigate the relationships that these other broader factors have with the risk of admission. So that was a policy methods decision we wanted to share with you and that we can discuss more.

**And finally, one of the wonderful things about working in this environment is we have many, many patients, so these ACOs have from hundreds to thousands, even more than 10,000 patients in the measure cohort. And you can imagine that's much different than what we see in the hospital-based measures. As result we are able to easily detect statistically significant differences.**

**And so, you'll see that as a group they show highly varied performance. On average they do better than fee-for-service providers on this outcome, and we norm the measure against the national group of fee-for-service providers that if they are doing better, they will show up as better. But the range of their performance is really wide. So for example, the minimum and the maximum risk-standardized acute admission rates is after risk-adjustment. For heart failure it ranges from 54 to 120 person-years, which is very a clinically policy-meaningful range. For diabetes it's from 24 to 68. For the multiple**

chronic condition measure it's from 48 to 107.

So finally, I just wanted to note we, as you know, submitted a supplemental memo on disparities analysis using the AHRQ nine-digit ZIP code. And in that more updated analysis, which it looks at 2013 ACOs, 220 ACOs, we still see -- we do see variation and somewhat of a trend with the ACOs with the largest proportion of low-socioeconomic status patients having slightly higher scores on the measure.

But even in the group, the quartile with the most low-SES patients, we see for each one of these measures that 30 to 40 percent of the ACOs are performing better than the national rate. We see some really outstanding performers, which is consistent with, for example, what other ACO Programs have seen. The one I'm thinking of right now is the Blue Cross Blue Shield Quality Contracts where some of their best performers have been provider groups that have a high burden of low-SES patients.

So it's a really exciting environment

to be working in and we're really looking forward to your input. I'm going to have Erica say a few words about heart failure before we focus on the measure. Thanks.

DR. SPATZ: Great. Thanks, Elizabeth.

My name is Erica Spatz. I'm a general cardiologist and I led the heart failure measure. And so, I'd just like to take a moment just to focus specifically about heart failure and how this measure we think helps to advance quality of care.

So the heart failure measure is a risk-standardized measure evaluating quality of care of heart failure patients cared for ACOs. In conceptualizing this measure we focused on hospital admissions, because for patients with heart failure hospitals admissions are associated with high morbidity and stress, they increase their risk for dying, and they're extremely costly to the healthcare system. So we think that we're measuring something that's important to patients as well as to healthcare systems.

And so, why is this measure all-cause unplanned hospitalizations needed? For people with heart failure they are at risk for a range of different kind of admissions. It's an extremely vulnerable population. So they're at risk for heart failure exacerbations, AFib and other cardiovascular diseases. They're also at risk for related complications like renal failure or electrolyte disturbances, but they're also at risk for a range of other hospitalizations.

For example, due to hemodynamic instability. They may be at risk for falls due to immune incompetence. They may be at risk for pneumonia. And we wanted to capture all of these different kinds of complications and really incentivize a range of providers, not just cardiologists like myself, but to work with a range of providers throughout the healthcare system as these ACOs are set up to do to provide hospitalizations in this very vulnerable population.

We wanted to move beyond the only

other outcome measure for patients with heart failure, which is the AHRQ PQI, the Prevention Quality Indicator. That PQI only measures heart failure exacerbations so it misses almost two-thirds of the other kinds of admissions that people with heart failure are coming in for. So we think that this measure adds to what currently exists.

I wanted to highlight also that we think that admissions are important because we have really good evidence that we can lower the risk of hospitalization by providing highly coordinated care, by providing care navigation for people, home-based services, participation in cardiac rehab. These are examples of interventions that have reduced hospitalizations in the heart failure population.

The third thing I'd like to highlight as you review our measure is the risk-adjustment model. We take into account a range of risk factors that increase the risk of admission, including a variable capturing pacemakers and

ICDs. We know that prior hospitalizations is a significant predictor of future hospitalizations, but we don't adjust for them. And the reason that we don't adjust for them is that we think that prior hospitalizations are a marker of quality of care. We don't want to adjust for variables that can confound quality. So that's one of the reasons why you won't see that in our risk-adjustment model.

As Elizabeth referred to as well and as you heard about yesterday, we also don't risk-adjust for socioeconomic status. And our team put a lot of thought into this. We know that these patients have a lot of challenges. We know that they're at increased risk for admission, but we also know that in particular the ACO Programs are designed to work with their patients, to work with their communities to mitigate the effects of poor access to care, medication non-adherence, health illiteracy, increasing opportunities for healthy eating, cardiac rehab. These are things that ACOs can actually improve for patients to

improve their outcomes.

So, and as Elizabeth mentioned, we found a number of positive deviants in the ACO population, meaning ACOs that were caring for the most number of low-SES patients, and we did a number of things to look and see that these patients were truly low-SES. We looked by the nine-digit AHRQ SES Index, we looked by Medicaid dual-eligibility to really clarify that these were low-SES patients. And in that group of ACOs that were providing care to the most number of low-SES patients there were a significant number, almost a third, that were performing better than the national rate.

So we are excited about these ACO programs because we think that they can meaningfully impact outcomes and we think that this measure helps to eliminate those differences and drive quality care improvements.

So we look forward to your comments and happy to answer any questions that you may have.

**CO-CHAIR BULGER: Thank you. So we'll go to -- the discussants are Bruce and Karen. And Paula was not able to be with us this morning. So we'll start with Karen.**

**PARTICIPANT: You want to go one by one?**

**CO-CHAIR BULGER: Yes, we're going to go one by one. We'll talk about evidence first.**

**MEMBER JOYNT: Yes, I think I agree with the measure developers that this is a measure that fills an important gap and has evidence, at least in terms of the theoretical relationship between clinical interventions and the ability to keep people out of the hospital.**

**At what point would the AHRQ PQI -- should I leave that alone for now, and the overlap with other admission measures or any differences with other admission measures? Should I leave that alone for now?**

**CO-CHAIR BULGER: Yes, we'll come back to it.**

**MEMBER JOYNT: Okay.**

MEMBER BRUCE HALL: I agree with Karen. So in general, as you heard, unplanned admissions per 100 patient-years with the diagnosis, fee-for-service Medicare data, all-cause admission except for the planned algorithm, LVADs excluded, transplants included, conceptual model seems very sound as they portray in their figure 1. So in terms of importance, literature provided seemed to support the importance of the topic.

CO-CHAIR TRAVIS: Any comments from the Committee members?

(No response.)

CO-CHAIR TRAVIS: Okay. Oh, go. Sorry, Leslie. Go ahead.

MEMBER LESLIE HALL: It seems in all three of these, as you're bringing up all three of these, there is overlap. And potentially an intervention for a patient with diabetes might be a heart-related intervention and a primary diagnosis might actually be any of all three of these. How do you account for the overlap and

the interventions that might be compatible with one measure, but actually affects multiple measures? How do you get around the confusion?

DR. DRYE: Okay. That's a great question. So another thing that we had to grapple with in this setting that is different from the hospital measures that are focused on admissions for an acute condition, there is overlap in the cohorts with these three measure and the patients included.

And so, by design a multiple chronic condition measure includes patients with heart failure. It doesn't have diabetes as a qualifying condition, but half the patients who have multiple chronic conditions have diabetes, so many patients there are also in the diabetes measure. And I can give you some numbers if this doesn't -- I'll just give you the amount of overlap and then I want to address your question.

So about a million patients have heart failure, but not by diabetes. This is in the Medicare fee-for-service cohort in 2012 that we

used for this measure. And about a million have only diabetes but not heart failure. And about 1.1 million patients have both conditions. And then those groups do overlap with a multiple chronic condition cohort in about 30 -- actually it's about 36 percent. So 1.8 eight million patients are in the multiple chronic condition cohort with neither of those conditions.

So, yes, there's a lot of overlap, but there's also a lot of distinct patients in each group. And I think it's okay. I mean, the measure performance is going to be correlated because it's some of the same patients. And some of the interventions will cross. I think that's programmatically okay to have them in those three measures, nonetheless in the same program, because it provides different information to the ACOs. But, yes, there is overlap.

MEMBER HEIDENREICH: Yes, maybe this is just for CMS, but I'd encourage them to coalesce around maybe one chronic condition measure. I think when we've looked at people on

heart failure in the VA who are identified off their outpatients; so you need your two diagnoses, a lot of those people you wouldn't think that was heart failure patients, their heart failure is number seven on a list of a ton of things, and it's a completely, or a very different cohort from those with discharge with a primary heart failure. And I think the interventions really aren't heart failure-specific that we're thinking they could benefit from.

So I would encourage you to go -- in future potentially combine these.

MS. SHAHAB: John, Frank Briggs had a comment as well.

Frank, did you want to --

(Simultaneous speaking.)

MEMBER BRIGGS: Good morning. I had a question. Since these measures involve the ACOs and the ACOs are coming and going into that program at different starting points, unlike hospital readmission programs and such where

everybody was measured starting essentially at the same point, the interventions given by the developers, while I believe can make a impact, many of them take a long time to establish and then to really see that impact, but how do you adjust for an ACO coming into the program, new into the program as compared to an ACO who might have been into the program for three, four, five years and may have been working and have these interventions established?

DR. DRYE: Good question. So one challenge for these measures versus again the hospital-based measures is that there isn't a clear time zero, or as clear of a time zero before which we want to adjust for the patient status and after which we're going to evaluate the outcome. It's a bit arbitrary. Like you say, there's action going on all the time and if you improved your patient last year, the risk factors will be lower. And so we won't estimate as high of an expected rate of admission for that patient.

But we tried to be pretty structured about that. We vetted it in public comment and with our expert panel and what we do is we just take the start of the measurement year as the beginning point for evaluation. We accumulate risk factors up to that point and then we evaluate for the outcome. And that is going to capture ACOs at different points in their progress.

I would just add that in the ACO Program they've just added another dimension, which is measuring year over year improvement in the quality measure set and allowing ACOs to earn bonus points for improvement. Separately from this project we're working on a method for that and for measuring that on risk-adjusted outcome measures specifically on these measures, which we'll share later at AcademyHealth later in the month. But there is not a clean start-stop time that we can identify. We would expect to see ACOs that are improving, getting better on this measure score and also that improvement showing

up in the improvement component of the ACO evaluation.

CO-CHAIR BULGER: Ye, I mean, I think these measures, just like most of the other ones, the crux of the matter comes down to how they're used. And from an ACO perspective the end-all, be-all is the total cost of care for the population and using measures such as this will help you drill into the total cost of care. But I think there is concern about how all the overlap is among all the measures.

And I think one of the things I do now is running our ACO and measures like this are very helpful again to improve the quality and to really drill into that total cost of care, I think. And we'll probably talk about this more when we get to use, too, but how they're used is going to be important because, this has already been said, they overlap with a whole bunch of the other measures as well.

Any other comments?

(No response.)

**CO-CHAIR BULGER: Okay. Vote on evidence.**

**MS. HERRING: Voting is now open on the evidence criterion for Measure 2886. Your choices are one, yes; two, no. We're looking for 18 votes.**

**(Voting.)**

**MS. HERRING: If everyone could submit their votes just one more time. Sometimes it takes a minute for the clickers to wake up since this is the first vote this morning.**

**(Voting.)**

**MS. HERRING: The results are 18 yes, 0 no, so 100 percent yes.**

**CO-CHAIR BULGER: Okay. So we'll start with Bruce this time and go with the next question on evidence.**

**MEMBER BRUCE HALL: Opportunity in terms of opportunity or gap. The developers quote a pretty impressive unplanned admissions rate of about 85 per 100 person-years in the crude Medicare population, fee-for-service**

population. And then in those assigned to ACOs about 82 per 100 person-years with the range, as they mentioned earlier, being about 53 to 121. So quite a bit of apparent variation in opportunity.

I terms of the ACO level scores, about half of the ACOs were rated as either high outliers or low outliers by their method with actually about twice as many being rated high as low as they portray in their figure 3. So a couple of axes indicating that there's opportunity for improvement.

MEMBER JOYNT: Yes, I agree. I think the gap is probably greater for admissions than readmissions, especially since you're moving -- if you move your admission denominator, you can change your readmission denominator and I think the evidence here would suggest both in terms of the ability to find outliers and probably in terms of the ability to focus attention that there's a gap here that's quite notable.

CO-CHAIR BULGER: Any comments from

the Committee?

(No response.)

CO-CHAIR TRAVIS: On the phone?

(No response.)

CO-CHAIR TRAVIS: Okay. So we'll vote on the gap.

MS. HERRING: Voting is now open for performance gap on Measure 2886. Your choices are one, high; two, moderate; three, low; four, insufficient. We're looking for 19 votes.

(Voting.)

MS. HERRING: The results are 11 high, 8 moderate, 0 low, 0 insufficient, so 58 percent high; 42 percent moderate.

CO-CHAIR BULGER: So to reliability, Karen?

MEMBER JOYNT: So reliability testing was performed at multiple levels, and the sort of test-retest as well as the data element reliability were good.

I have a couple clarifying questions, if that's okay.

The reliability, I couldn't tell from the documentation if it was in the total sample or in the ACO sample. Particularly with some of the new programs coming out of CMMI and CMS for support trying to get smaller ACOs in and more rural ACOs, is there a limit at which the reliability starts to fall off by size or are they all big enough that that wasn't an issue? I couldn't tell what the locus was of reliability testing.

DR. DRYE: Well, as I mentioned before, the smallest ACOs are a couple hundred patients in each one of these cohorts. For diabetes it's even bigger than that.

The way we tested it was we include in the measure score calculation all the fee-for-service patients and then we compared just the -- we split the sample, we compared the risk-standardized acute admission rates from those two random samples from each ACO, just like in the measures we discussed yesterday, and we looked at the ICC, which here is 0.81 or above for all

these measures.

We don't do that by looking individual at each ACO per se. I think that might be what you're asking, Karen. I'm not sure.

MEMBER JOYNT: Well, for the hospital measures you have a 3-year sample and generally under 25 doesn't get --

(Simultaneous speaking.)

DR. DRYE: Yes.

MEMBER JOYNT: I wasn't sure if that had been empirically derived and if so if there was an empirical derivation here.

DR. DRYE: Oh, for the minimum? Yes.

MEMBER JOYNT: Yes, whether it's reliable, or if that's sort of a future --

(Simultaneous speaking.)

DR. DRYE: Yes, so that's a great question. The measures are -- it would be great that -- see, the contractor working with CMS that's actually crunching the numbers for the 2015 data might be able to help us here. I don't think we're going to see any of these ACOs have

an unreliable score. I don't think they've had to set a minimum. I could say some of the learning around that really came when we started working on measuring improvement year over year in the year after -- in this past year, because we could actually fit a whole model right on every single ACO. There wasn't a single ACO that didn't have enough data for us to fit a GLM easily. So I don't think there's any ACO that is too small.

The minimum ACO size in the Medicare Shared Savings Program is 5,000, but these are such common conditions that we've been having absolutely no problems with sample size. But I can get back to you specifically on that just to see if there was anything that was even close to too low.

CO-CHAIR BULGER: Actually, I do want to follow up on that because while we're talking about the measure for ACOs -- it could be used for any cohort, correct? So let's say for example a state wanted to go into a global

payment program like Maryland does and wanted to assign hospitals a -- each hospital has a cohort of patients. They could do that and they could end up having lower numbers of patients. I mean, the current Medicare Shared Savings Program and other ACOs do have minimum numbers. And it wouldn't get to the number, but there is a possibility that someone -- if this is endorsed, someone could pull this measure and use it for smaller cohorts than a current ACO Program has.

DR. DRYE: I agree. And I just wanted to ask if our statistician Haikun Bao at CQN from Yale is on the line. Do you want to make any comments?

DR. BAO: This is Haikun. I think we -- for the ACO we have a -- for individual ACO we have enough for sample size. So for each individual ACO I think that this measure is okay.

DR. DRYE: And I think just to add also; and Jeff Herrin who's one of our statisticians on the hospital measures is here today, when we got that number 25 on the hospital

side, we actually did a lot of modeling to figure out what kind of -- what do we need to get a stable estimate. We just haven't done that for these measures because we didn't need to, but if you were going to apply it in ACOs with fewer patients or some kind of population, health plans that might be small, yes, you would need to do that testing.

MEMBER LIND: Okay. So given the shift from inpatient to outpatient and the impact that people think that the readmission penalties are having, notwithstanding some articles to the contrary, did you look at what effect you might have in terms of outcomes if you included the outpatient set, or at least the emergency room contacts? I mean, why would you restrict it to just inpatient admissions, I'm just wondering, since presumably there's -- in terms of the quality of ambulatory case care-sensitive management, condition management, it seems like an emergency room contact would be as important? Maybe not quite as important, but it would be

important.

DR. DRYE: Your last words are exactly the reason. I mean, we were looking for an indicator of acute decomposition, and if we moved to ED visits or ob stays, it just -- there's different -- as you know, different providers have that bar in a different place. When would you admit versus treat in the ED? But we just felt like using acute unplanned admissions was a pretty high bar and relatively more even across providers. But there isn't a reason you couldn't evolve towards looking at ED visits, too.

And as I say, for CMS under this same contract to build outpatient outcome measures we are looking at -- we combine actually ED visits, observation stays and hospitalizations for post -- a measure of colonoscopy quality, because the predominant outcome there is ED visits, not admissions. So we looked at that, but for this measure we wanted to keep the bar pretty high.

I don't know if you want to add anything for heart failure, Erica.

DR. SPATZ: Yes, I think I would agree with that, and I think that that's somewhat validated by the high proportion of admissions that we see in these very vulnerable groups where they are meeting a threshold for admission.

Lynne Stevenson, one of the giants of heart failure who talks a lot about this, really just presents a model which is kind of interesting to think about because we're so focused on heart failure readmissions. And she talks about these three phases of heart failure management, and one is the transitional management that is right on the heels of a hospitalization where we're talking about high touch and all the interventions that we're testing to lower readmissions.

And then she talks about this plateau phase. And it is complex care. It's multidimensional care, it's looking at their medications, looking for both evidence, guideline concordant medications, as well as drug interactions and potential adverse events. And

we think that the outcome of admissions is reflecting that high vulnerability. Certainly ED admissions are important and might be something to consider, but in this very vulnerable population the hospitalizations alone kind of stand on their own.

MEMBER LIND: I mean, I think I see your point with congestive heart failure, but I'm thinking on the other two cases also where diabetes, maybe hypoglycemia, ketoacidosis, maybe a leg ulcer could be managed in the ER. Multiple chronic conditions could be -- not meet that inpatient threshold. It seemed like it might be more likely, might be more useful to have more information in those less-acute situations.

CO-CHAIR BULGER: Leslie?

MEMBER LESLIE HALL: Yes, I just had a question about -- as we transition from fee-for-service to any sort of value-based bundle payment or any new payment does that put anything at risk as far as data collection, because your data collection is now coming from fee-for-

service-only claims. And so, how do we evolve something that is touching such a broad group of people when the data collection methodology may actually be incented to be pulled away from the data elements you're requesting?

DR. DRYE: Helen, how are we going to do this?

I agree that's a threat over the longer term to the data that we're using. And right now Medicare fee-for-service is our one full national data set that we can use for these kind of analyses and has high feasibility. I think we have to keep an eye on that.

My understanding of the -- within the ACO Program all the claims are still filed, and that's partly how their shared savings is reconciled, but I think it's a dynamic situation that as -- those of us working in quality measure, we have to keep thinking how can we get the right data to understand utilization as these incentives for actually filing a claim are diminishing.

CO-CHAIR BULGER: Cristie?

CO-CHAIR TRAVIS: Well, I couldn't help but think about Keith's question and then remember the measures that we looked at yesterday for excess acute care days for heart failure, specifically on the hospital side.

So I think to a certain extent -- and maybe I'm not thinking about this correctly, but in the ACO environment there are -- I guess are there hospitals? And I would think there are. And then so, to a certain extent if we move to looking at holding hospitals accountable, the excess days which include ER and observation as well as inpatient admissions, then we're beginning to move into that arena for heart failure anyway.

And so, I guess I'm trying to kind of reconcile the discussion we just had with the fact that we looked at those measures yesterday, albeit for a different -- it wasn't for ACOs, but they're usually part of ACOs, I would think.

DR. DRYE: Yes, great. And great

point. And some of these outcomes are overlapping. The ACOs -- I mean, this just really surprises me, but fewer than half actually have a hospital, but many of them do and most of them don't. So we have over 400 ACOs right now. And so there are incentives to hold down costs. So that's one difference.

And the other difference is the cohort. So those hospital-based measures that are looking at patients who have been hospitalized for heart failure, we're looking at anyone that is diagnosed with heart failure. So there's overlap. It does mean that when you look at these scores they travel together to some degree, but I think that's one thing we could spend more time really understanding is how they're traveling together. but they are looking at separate domains of quality and at different -- they're profiling different providers, even though there's some overlap in the providers as well.

CO-CHAIR TRAVIS: Wes?

MEMBER FIELDS: Sorry. I want to follow this line for a second just so I understand your methods and how the data is going to flow.

One of my theses is that there's a lot of really creative things going on inside of the Medicare Advantage groups and the integrated systems and that it's for a number of reasons obscured in proprietary data. And there's reasons why it doesn't get shared with CMS.

But my question is CMS did provide a waiver on the three-day rule, for example, to Pioneer ACOs. I believe that was in '14, calendar '14. And I've tried to track this at the local ACO level with a fairly large one in Southern California, and it's hard to get at the data. So I'm just curious, in terms of episodes of care other than admission are you tracking admission to short-stay SNF facilities, for example?

DR. DRYE: Okay. Let me just answer that last part, and then I'm not sure if I fully

-- if you wanted me to address part of what came before.

So in terms of SNFs, we aren't looking at it. I know in the hospital setting we've looked at whether readmission rates are related to the prevalence of SNFs. The reason we're not focused on SNFs per se is the use of SNFs is so variable across the country and it's definitely going to vary widely across ACOs. So we try to just be neutral with respect to those kinds of factors that if we use them they're going to perturb our scores. So we're ignoring SNFs per se. And so, is your concern that maybe that people are diverting to SNFs or -- I'm not --

MEMBER FIELDS: No, nothing as nefarious as that. I've heard an ex-CMS administrator refer to ACOs as HMO-like, and one of the reasons I think that's true is, at least for the acute care continuum, one of the things which is pretty easy to do without the statutory requirement for a three-day inpatient stay on the Medicare Advantage side is to be able to

aggressively evaluate and refer somebody with MCC, especially things like diabetes or a true MCC patient, to basically preemptively screen, stabilize and transfer to a skilled nursing facility in lieu of what would be a hospitalization for a traditional Medicare A/B patient.

And so, my point is that CMS has been enlightened about providing a three-day waiver to ACOs, but that means I think for you to really track innovation well you really need to -- I think you should reconsider the inclusion of SNF services.

DR. DRYE: I don't know, also someone from CMS may want to comment on this if they're on the line, but I think it's a really interesting point. I mean, Susannah Bernheim mentioned yesterday we do a tracking of some of these effects of measurement for CMS as part of our broad set of work. And I think earlier on we were a little more focused on SNFs and thinking about how they might be related to readmission.

I think it's a good question.

My advice or my suggestion would be to look at that, but I don't think we would add them into the measure, again because if you start adding SNFs into a measure -- SNFs are just -- could be SNFs versus home health versus some other strategy. So we try not to pull provider types that -- post-acute care choices into these types of measures.

CO-CHAIR BULGER: Derek and then --

MEMBER LIND: I think the three-day waiver only applies to two-sided risk ACOs, not the one-sided risk. So it's only the dozen or so Pioneers.

CO-CHAIR BULGER: Correct. Derek?

MEMBER ROBINSON: Thank you. The discussion regarding the migration of your data source as we move from claims-based payments to other models just sparked another question in my mind, and that is that there have been a lot of efforts especially for activities focused on ACOs to try to harmonize the use of minimum measure

sets. And because it appears that this measure set is focused on the Medicare population, I just wonder is there any thought to the applicability of this measure to maybe an age group of 45 to 64, for example. I just say that because of some of the recent activity with the Core Measures Collaborative.

And so if you've got multiple payers designing a minimum core set of ACO measures and CMS likes this one and says, hey, this is a great measure to move forward with but then it's not applicable to other age groups, is there a way to harmonize that? So just a question.

DR. SPATZ: Thanks. It is a good question, and we are always considering who else is this applicable to? I think in the case of heart failure we need to pay careful attention to our ability to adequately risk-adjust, especially in the younger population where there are centers of excellence that young people with cardiomyopathies or advanced heart failure will go to. It's a little bit less of a bread and

butter population than in the elderly population, which is not to say that we couldn't go to non-Medicare patients.

But I think it's a really important consideration that needs a lot of thought and development to make sure that we can adequately risk-standardize across ACOs so that we're not getting discrepancies in performance based on differences in case mix.

DR. DRYE: I would just add that one of the things, one of the reasons it's a privilege to work on these kind of measures is we are very transparent with our methodology. We will share it with anyone who's interested, whether that's a state trying to move under the state innovation models to a core set of measures or Blue Cross Blue Shield or others who are looking to set up -- use measures in their all-payer environment.

So you can ask for status specs, you can ask for the specs. And we have done testing of many measures in the all-payer setting and

looked for how well do they stand up in that 18 and over population. We haven't done it with these measures.

CO-CHAIR BULGER: Other questions from the Committee?

MEMBER JOYNT: Sorry, I was just flipping through this. So this is 65 and older only, right? It does strike me that -- I agree that under 65s may be different, but they're also really important, really expensive and really under-studied. So the rationale is just that you can't risk-adjust well enough for the under 65s, or they're different somehow?

DR. DRYE: Yes. Well, for this measure we haven't pulled in those that are eligible for Medicare who are under 65 because they're typically a lot sicker. And in all of our measures that we're using fee-for-service Medicare data for we made that decision.

But I'm totally sympathetic. I think we'd want to be able to in a lot of settings -- for example, in in-state innovation models where

we're trying to come to single core sets of outcome measures that are population-based, to bring that all the way down to whatever threshold makes sense. I mean, for COPD we'd bring it down to 40 in some settings.

So it's just these are new and we haven't gotten there yet. And I think in the program, the Medicare Shared Savings Program, this is -- it makes more sense to keep it at 65 and older.

DR. SPATZ: Just to clarify, the under-65 group that are Medicare fee-for-service are a really unique population, because to qualify for Medicare under 65 -- there's very special populations that do so. And so, that does raise their level of severity of care. End-stage renal disease patients, for example.

MEMBER JOYNT: I think many have argued that's exactly who we should be trying to prevent admissions in, right? I mean, and if you look at the MedPAC data for Medicare Advantage, and actually claims data for -- encounter data

for Medicare Advantage are now available. So in theory we could be rolling this across multiple types of patients. But anyways --

DR. DRYE: Well, just to that point, if I could, we would love to have the data for Medicare Advantage and pull those in. That's almost 40 percent of the --

MEMBER JOYNT: Right.

DR. DRYE: Right. So that -- we're working on that.

MEMBER JOYNT: Yes, I know. Speaking to that converted on the data issue.

But with the under-65s MedPAC has shown, I think, pretty convincingly in the MA data that disability is a bigger driver of core outcomes than dual status. And if you're looking at really the vulnerable population, preventing admissions, running a good ACO and saving money, that may be where you need to go. If you feel methodologically that group can't be included because they just are so different, it may be helpful to know why that's the case, because I'm

not sure that excluding the 65 and under from all fee-for-service measures is a good long-term rule.

DR. DRYE: I agree with you. I mean, I think it's just another -- it requires further work and evaluation. I think we will be going in that direction.

CO-CHAIR BULGER: Larry?

MEMBER GLANCE: Do you have any concerns of the potential for unintended consequences with this measure? It may be that too few admissions is a bad thing for people with heart failure and may lead to excess mortality. This is a little bit new for us in terms of quality measures. But in obstetrics, for example with C-sections, driving the C-section rate down too low may be a bad thing. There may be more of a kind of a U-shaped curve. And rewarding the hospitals with -- the ACOs with the fewest hospital admissions may not actually be the right way to go here. I don't know if you've thought about that.

DR. DRYE: Yes. And again, I think we're flipping back -- between talking about heart failure to really talking about all three of these measures. Just realizing we're talking about -- so if we're not confusing you.

They're part of a broad set of ACO measures. So I'm just looking at the measures that are changing next year -- well, for this year, 2016-2017 that include multiple domains of care including patients' experience of care, so the CAHPS surveys. And I think that's really critical. And then we have to be looking at a broad set of outcomes when we use these measures so that we're understanding health outcomes, patients' experience, as well as cost.

It's a question that applies to the ACO Program or any shared savings program overall. Are we creating incentives to provide too little care? And I think we have to be vigilant about that. We tried to build tracking of some of those effects into what we're doing with CMS, but I think more broadly CMS is looking

at that as well. And some of the other measures in the ACO set cover those things, but not perfectly.

CO-CHAIR BULGER: Karen?

MEMBER JOYNT: Yes, just a quick response to that. The quality measures in an ACO have so very little impact on the money they save compared to the actual money they save that you could argue that being in an ACO makes you want to cut admissions much, much more than this measure does. I've certainly heard it said convincingly by some health economists that having admission and readmission measures in an ACO is completely redundant. You already have very strong incentives in place to cut admissions. That's where your dollars are.

It's a separate question, but it does a bit beg the question of evaluating a measure for its use within a program versus evaluating the measure for being methodologically sound for the population to which it's applied, which I think it is. The use I think is actually bigger

and more complex question even than just this measure.

CO-CHAIR BULGER: Any other questions, comments from the Committee on reliability?

(No response.)

CO-CHAIR BULGER: Anyone on the phone?

(No response.)

CO-CHAIR BULGER: Okay. So we will vote on reliability.

MS. HERRING: Voting is now open for reliability for Measure 2886. Your choices are one, high; two, moderate; three, low; or four, insufficient.

(Voting.)

MS. HERRING: I believe we're looking for 20 votes this time around, so if everyone could just vote one more time?

(Voting.)

MS. HERRING: The results are 2 high, 18 moderate, 0 low, 0 insufficient, so 10 percent high, 90 percent moderate.

CO-CHAIR BULGER: Okay. Validity?

**MEMBER BRUCE HALL: Obviously the discussion of reliability ranged on a variety of issues including some modeling issues, but in terms of just reminding the Committee about some of the details of the measure that we've heard, this is a standardized risk ratio where the predicted and the expected are used to create a ratio that's then multiplied by the grand mean.**

**The methodology uses two years of data prior to the beginning of the performance period to define the diagnosis. The diagnosis of heart failure is defined by either one, inpatient, or two, outpatient codes in those prior to years. In the first year preceding about 90 percent of the patients are identified and in the second year preceding about 10 percent of additional patients are identified. And rheumatic failure was included as per expert advice. This is depicted in the figure 1 by the developers.**

**I did note that the program on the whole seems to be fee-for-service data-dependent, but I think we've already covered that in the**

last few minutes.

In comparison to the two years used to define the diagnosis, one year of data is used to risk-adjust the diagnosis, and that seems appropriate. And then the performance period is one year beginning on the beginning of the calendar year.

As a note, about 114 of the MSSP ACOs were included in the diagnosis, and as I noted earlier, about 61 of them were rated as no different than average with 37 being rated better, and 16 were so. About half of all the participants at the measure level are being rated as either better or worse and then two-thirds of those are being rated better and one-third worse.

For that 114 MSSP population there were roughly 120,000 patients. And so going back to Karen and some of John's comments earlier, we're probably looking at a mean patients of about 1,000 per ranging down to a couple hundred in terms of minimum necessary for future considerations.

The developers did consider SES factors. As we've heard both this morning and in more detail yesterday, they examined both the ACS-based AHRQ Index and dual-eligibility status. And I will say that while the developers continued to feel that the effect was small, and I support that, still from some perspectives there does seem to be a substantial effect. So for instance, if you just compare the top and bottom quartile risk by either of these, there does seem to be an effect.

So for instance, if you just look at the SES Index version, the fourth quartile has about a 25 percent admission risk while the first quartile has just a 7 percent admission risk. And if you go to dual-eligibility status instead, it's 24 percent for the worst quartile versus 3 percent for the best quartile.

So obviously those are crude figures, but I'm still concerned that there might be a substantial effect from some perspectives, understanding that at present the decision is not

to adjust further for gender, race, resources, behaviors or other aspects of the conceptual model that we want the ACO to be responsible for. We want the ACO responsible for encouraging good behaviors and so on and so forth.

And so I think it's conceptually sound that these things are not currently adjusted, but I do wonder whether some perspectives are showing us a pretty substantial effect.

In terms of other modeling decisions, any days in the hospital are subtracted from the out-of-hospital days risk. And then if a patient starts the year in inpatient and dies, they're totally excluded from the model because they don't contribute to any outpatient days of risk.

The variable selection for the model was driven by AIC, Akaike Information Criterion. That seems to be a very sound approach, the way the developers applied. And in terms of fit, about 12.2 percent of the variance is explained by the risk-adjustment model, which in terms of medical practitioners might not seem a lot to us,

but in terms of economists it's actually a substantial amount of variance to explain. Otherwise, the fit criteria appeared to be good as portrayed in figures 5 and 9 through 11. And the quartiles of risk at the end seem to be well estimated and fit pretty well.

I also had noted that under the current paradigm ACS can be entering measurement at different points of maturity, but I felt that that was very reasonable, because if ACOs, quote, "need to catch up," then this is the way to show them that they need to catch up. While at the same time, as Liz and the developers mentioned, the risk-adjustment scheme, because of the way the risk-adjustment is determined off the prior year's data, it does blunt that a new ACO is going to look horrible. It does blunt that, but it still I think preserves the ability to show an ACO that they have catching up to do.

So those were the concerns and issues I raised in reviewing. I think each is appropriately handled by the developers.

CO-CHAIR BULGER: Karen?

MEMBER JOYNT: That was very thorough. I only have a few things to add. One just kind of, I think in MSSP there's a hold-harmless where you get a year of reporting. I believe that any new measure everyone gets a free year.

DR. DRYE: Yes, two years of --

MEMBER JOYNT: Oh, two years?

DR. DRYE: -- pay-for-reporting and performance score doesn't matter.

MEMBER JOYNT: Oh, I thought there was only one. Anyway, so you do have an opportunity to see your performance before you're paid on it for these, but I had one question about the comorbidities, and I think this is more important for an admission measure than a readmission measure because you're trying to look at aggregate risk.

My understanding is that a lot of comorbidities in the fee-for-service population, much more so than MA, will fall out because things like quadriplegia are actually not

terribly stable in the Medicare fee-for-service data. I don't know if this is totally like an old wives' tale, but that because in Medicare Advantage comorbidities can be picked up by a set of not just claims-based ascertainment, whereas in fee-for-service there is only claims-based ascertainment that chronic disease that should not vary over time does vary a lot in claims.

And so I've heard it argued that for better ascertainment of chronic disease that a two-year look-back is favorable to one, but I'd just be curious if you did any looking to see for this type of stuff whether things actually -- it made a difference how far you looked back for chronic risk-assessment.

DR. DRYE: Yes, I mean, I think in this setting, and I just want to give a shout-out to our analysts, we're using so much data. We're looking at all the outpatient data over two years. And I don't think that we're losing things. We did look at -- we did think about going back three years, which is more burdensome

given the tremendous volume of data you have to process to get these risk factors for these very large cohorts. And we looked at -- for example, we were particularly focused on dementia and looking back, and it made a very small marginal difference there.

But I would say the one area we were the most concerned that we would not be fully capturing risk factors is dementia and mental health because they're often just not recorded by providers, raised by patients. So it's not going to be perfect, but I think we have a really broad sweep and a chance -- we were able to pull those from both inpatient and outpatient claims over a long period of time.

CO-CHAIR BULGER: Paul?

MEMBER HEIDENREICH: Just one thing. Since the SES was brought up -- and I suppose that goes to usability potentially more, but I feel there's a much stronger case for using SES than with the measures presented yesterday looking at the dual-eligible. And you have an

eightfold difference between quartile one and quartile four being in the worse than national rate of 24 versus 3 percent. I realize it's relatively small numbers, but I can't believe that that is all due to quality differences that we need to pay attention to.

And again, I think it's -- I'm not sure we're the ones to make these decisions. It would be nice if this came from a higher level. But I think this is clearly different from yesterday's data where the differences were a lot smaller.

DR. SPATZ: Right. So I think we struggled a lot looking at our data and how ACOs perform in the different quartiles of proportions of low-SES patients. There is a trend, and as you point out there are more ACOs who are poor-performing ACOs in that group that are caring for a lot of low-SES patients. So that trend does exist.

We also saw a lot of heterogeneity, which actually the heterogeneity increased. The data that we submitted were from 2012. The 2013

data show that even more ACOs that are caring for a lot of these patients are performing well.

When we think about is it fair to compare ACOs that are caring for different proportions of low-SES patients, no matter how you feel about it, when we look at our risk-adjustment models, we don't really see a big difference. So it does pose a challenge for risk-adjustment, because if we want to meaningfully consider the challenges of ACOs that are caring for these patients, the risk-adjustment models don't end up helping them. They're pretty bland.

DR. DRYE: I would just add, going back to the conceptual model, if you want to pull that back up, which is in 2b4.3 of the testing form, there are many factors in the admission context that we're dealing with that probably aren't as relevant even in a hospital setting like resources in the community, transportation, the physical environment, health behavior norms, the conversations that we've had the ACO Program

going back to is it a policy? It's a policy methods question, really.

We know from what we're learning about ACOs that a lot of them are creatively trying to mitigate these factors. CMS is actively setting up a program to investigate how can ACOs partner and other providers partner with conversation public health and social services to mitigate these factors. So even though we know that trend is there, the main two reasons we're recommending against adjusting one is there are many, many good performers with a lot of low-SES patients. So we feel like that should set the benchmark.

The benchmark for these ACOs caring for low-SES patients is good performance, and that's established. So we have 30 to 40 percent of these fourth quartile doing really well, doing better than the national average. We want that to be visible. And other ACOs are succeeding in effectively caring for these populations, we want that to be really visible in the measure.

And then it's not as high stakes as in

the hospital setting if there's two years of pre-pay-for-performance. And it's one of many measures in a domain. And also, you can offset any reduction in your shared savings portion allocated by improvement. So there are factors that mitigate. It's programmatic. We are evaluating in the context of this particular Medicare Shared Savings Program and on balance it seems most important to reveal these differences and learn from them rather than to adjust.

CO-CHAIR BULGER: And I think the other thing you said though is you haven't necessarily said that SDS doesn't play a role in this. What you said is that your risk models don't speak to it. So there may be other things that down the road you could adjust for which would level out the data, but when you went -- the risk model you were using didn't fix it.

DR. DRYE: Well, we used the AHRQ SES Index and dual-eligibility, and those made very little difference. But as you can see on this conceptual model, there are many, many kinds of

-- like we could adjust for health behaviors in the community and lots of other things that we know are related to admission risk at the population level. But we're working to try to understand how those factors affect rates and what this -- where providers are succeeding working to lower admission rates in spite of those factors.

So we want the measure to be revealing success there. I mean, over time I think, depending on how providers do, you could revisit the balance of that decision, that that's what we're recommending at the outset of these measures.

CO-CHAIR TRAVIS: Wes?

MEMBER FIELDS: Yes, a question that's partly about validity and partly about structure, so a bit of a relapse, I guess, but I'm wanting to understand. The analogy is a little bit like very large hospitals versus critical access hospitals. So are you not going to distinguish between categories of ACOs?

For example, to me there would be a big difference between -- even if they had a substantial claims-based or 1,000 patients they're taking care of, to me an ACO that's community-oriented that doesn't have a hospital partner and is highly reliant on community resources for all the things that Medicare doesn't pay for to keep folks out of the hospital -- to me that's a very different model than a large scale ACO piggybacking on a well-funded multi-specialty group that has other access to capital.

So the short question is, are you distinguishing between categories of ACOs or are you -- you're going to report on one pool of ACOs even though they have very different attributes?

DR. DRYE: Correct. So I think this is just going to be an area of learning for the field is further -- and with great questions about what types of ACOs are really doing well in these measures? But the ACO measure set CMS uses applies across the Shared Savings Program ACOs,

which, Karen mentioned, they don't have that much of a return on their investment or that much risk, and also the Pioneer and evolving next generation ACOs that are highly, highly varied in how they're structured. But we use one measure set.

CO-CHAIR BULGER: Yes, I think the other thing that comes up with that; validity I'm thinking about, is that comparison group is a self-selected comparison group, so it's not as if you -- the other measures that we're looking at, they looked at every hospital in the country or every -- I mean, this is a group that decided they wanted to do this and is a big piece of the --

DR. DRYE: Yes, that's a critical point. I want to clarify. You mentioned it before, but it's -- we actually -- we run the numbers against all fee-for-service providers, not just providers in ACOs, because we do not want the comparison to be just within ACOs. So we use hierarchical modeling like we do in a

hospital. We have so many cases that we get this great spread of performance, but the comparator is the admission rate among all fee-for-service providers. So in the analysis we have 20-plus million patients and then we calculate the scores just for the ACOs.

And so, you see on average -- because at the outset we didn't know what we were going to see, but we wanted to be able to see on average are these ACOs -- how they compare to the broad group of fee-for-service providers. We don't take it as a given, but they're better. But they did turn out to be, you know, have lower risk-adjusted rates. And that's how the program calculates it. When they use it in -- they keep all those other fee-for-service providers in the score calculation.

CO-CHAIR BULGER: Leslie?

MEMBER LESLIE HALL: I had a question about the admission data. Did you do this on ICD-9 or ICD-10 data, or both?

DR. DRYE: That is a very alive issue.

So the first performance measurement year is 2015, which has one quarter of ICD-10 data. So we have specified the measures in both ICD-9 and ICD-10. And we will be continuing to test the ICD-10 specifications as more data rolls in working with CMS and RTI, which is the CMS contractor that's actually crunching the numbers to produce the measure scores.

MEMBER LESLIE HALL: Were the gaps materially different between ICD-10 outcomes reported and ICD-9?

DR. DRYE: I don't think we know. We haven't seen that ICD-10 data yet. We'll be looking at it later this summer. What CMS is working to do is really look for -- looking across those four quarters. So if the fourth quarter, October to December of 2015, is coded ICD-10, then that data will be part of what's used to calculate the measure score. So doing quality checks to make sure that it looks like it should look. But it's a transition that we'll just -- we'll be testing through into the next

couple of years, I think.

As Karen Dorsey didn't say yesterday to you, but she was reminding me and said it's a couple of years before we're really, really confident about ICD-10 data. It's a big focus of our efforts right now to ensure the measures continue to be valid as we go through the transition. I'm not sure that was reassuring, but that's where we are.

CO-CHAIR BULGER: Kathy?

MEMBER AUGER: I have a methods question. So because the identification of the CHF population relies on two years of claims data and then your comorbidities rely on one year of claims data, if you're 65 and just coming into it, there's a good chunk of that population that just doesn't have claims data because they weren't in Medicare before. So I assume that those are at high risk for misclassification, that they may actually have CHF and you just can't recognize it.

What about also the patients who were

previously in Medicare because they were disabled? Are those previous claims like when from they were 64 used, or how does -- does this measure effectively -- a measure of patients that are 67 and older, so where they all have two years of claims?

DR. SPATZ: And, Elizabeth, feel free to jump. We used two years of data because we wanted to capture the healthy heart failure population. So we require two claims-based encounters with heart failure, if those claims are in the outpatient setting, only one with a principal discharge diagnosis from the hospital. However, we only require that people in the cohort have one year of prior data. That captures the majority of our cohort.

And we also thought that we needed one year of data to adequately risk-adjust for this population. Where to draw time zero was a question of when we start to hold providers accountable. And kind of consistent with prior measures we used the one-year time point.

With regards to your question of -- we are capturing the incoming population of people who are 65 -- is that correct? Do you want to --

DR. DRYE: Yes, it is a limitation in that if they were not enrolled as a dual-eligible, we won't capture them. If they were, we have their data and we will capture them in that first year of enrollment.

CO-CHAIR BULGER: Okay. Bruce?

MEMBER BRUCE HALL: I was just going to reiterate, add a comment back to the SES conversation again as pointed out by Liz and as asked previously by Karen. And that is that we are talking about a population where the payment policy, the penalty is already in place, so we're talking about designing a measure, as Liz said, to shed light on what's happening. Even if ACOs are advantaged or disadvantaged by their population they're under the payment penalty, so to speak, regardless already. And so, I think as I was reviewing the measure I felt it was a little bit of additional justification to take

the approach of wanting to shed light where there are issues.

CO-CHAIR BULGER: Other comments on validity? On the phone?

MEMBER BRIGGS: No.

CO-CHAIR BULGER: Keith?

MEMBER LIND: I'm sorry. I didn't understand your comment, Bruce, about what penalty we're talking about.

MEMBER BRUCE HALL: Well, by definition they're under financial constraint. As an ACO they have risk already.

MEMBER LIND: But not the one-sided risk.

MEMBER BRUCE HALL: Okay. Fair enough.

MEMBER LIND: The Medicare Shared Savings is one-sided only.

MEMBER BRUCE HALL: Yes, fair enough.

MEMBER LIND: There's only a half a dozen or a dozen that have two-sided risk, I think.

MEMBER BRUCE HALL: Right. Fair enough. Good comment. Under ultimate state, yes.

MEMBER LIND: And just to clarify the -- I don't know how relevant this is, but the readmission penalties I believe are waived for two-sided risk ACOs, not for the one-sided risk.

CO-CHAIR BULGER: They are not for the one-sided risk, I know that.

MEMBER ROBINSON: And I guess to tack onto that comment, moving forward with the APMs in the future under MACRA would be two-sided risks.

CO-CHAIR BULGER: Maybe.

MEMBER ROBINSON: Okay.

MEMBER HEIDENREICH: Just a quick -- I strongly agree we do need to shed light on this, however, a performance measure means ready for public reporting and labeling ACOs as delivering bad care. And I think that's where I'm struggling.

CO-CHAIR BULGER: Okay. So we're

going to vote on validity.

MS. O'ROURKE: Sure. Before we vote, I did want to just make sure everyone is aware that per the validity algorithm moderate is the highest level this measure would be able to get to due to face validity. So that is why you don't have high as a voting option.

DR. DRYE: Sorry. I just want to clarify one thing about the labeling, because the data we provided, the analysis we provided for this application classifies ACOs categorically into better, worse or no different. We have a lot of statistical significance and we have so many cases, but that's not the approach the ACO program is using. They're actually just creating a scale of -- a range of scores and setting a benchmark and a threshold. So they don't propose and they won't be labeling ACOs as better or worse.

So for both the better and the worse they may just be marginally -- we have small confidence intervals because we have so much

sample size. They may just be slightly above or below the national mean. But that's not going to be visible in public reporting as CMS is carrying it out.

CO-CHAIR BULGER: But it could be.

DR. DRYE: If they move to this kind of categorical approach, they could. That's not what they do in the ACO Program though. We used it --

CO-CHAIR BULGER: No, I understand but --

DR. DRYE: Yes.

MEMBER HEIDENREICH: That's my point. We're not judging how CMS is going to use this. We're going to say for anyone out there this is a way to report the quality of ACOs. And that's the concern. I agree CMS is doing it potentially the right way, but I'm not sure everyone will.

CO-CHAIR BULGER: Okay.

MS. HERRING: Voting is now open for validity for Measure 2886. Your choices are one moderate, two low, three insufficient. And I

believe we're looking for 20 votes. Oh, there we go.

The results are 14 moderate, 6 low, 0 insufficient, so 70 percent moderate, 30 low.

MS. WATT: All right. We'll go to use, which we've talked about a lot already.

Karen?

MEMBER JOYNT: Would it be appropriate to bring up the PQIs now? So there are competing -- I shouldn't use that term. There are other measures that look at admissions, the AHRQ Prevention Quality Indicators. They're obviously different. The big difference is that those are only risk-adjusted for age and gender category, which is because they were developed for use at a large geographic area. They're now being used -- I don't understand the whole endorsement thing. They're now being used for physicians, which is a not very large geographic area. And I believe they're in the ACO Program also, again not risk-adjusted.

My understanding is that AHRQ is

considering what to do with those measures and if they should be risk-adjusted, but they are very similar and overlap. The age range is not quite the same. They do include the under-65s in those, which you can decide whether or not you think that's good or bad. I guess COPD/asthma has different cutoffs. They're different. There's a lot of them and they have acute and chronic composites. They look at a very similar thing, which is admissions to the hospital. In their case for heart failure or for urinary tract infection or for pneumonia.

So they're selected based on why you were admitted, given that you are theoretically eligible. All three measures, to my understanding, are admitted for anything, given that your outpatient diagnosis has classified you for something. I don't know if we're supposed to consider that now, but they are very similar measures.

MEMBER BRUCE HALL: John, and we're on feasibility, right? Use is next?

MEMBER JOYNT: Oh, sorry. He said usability. It's very feasible.

MEMBER BRUCE HALL: No, all comments still pertinent when we vote.

CO-CHAIR BULGER: Any comments on feasibility? All right.

MS. HERRING: Voting is now open for feasibility for Measure 2886. Your choices are one high, two moderate, three low, or four insufficient.

And we're just waiting on three more.

The results are 12 high, 7 moderate, 0 low, 0 insufficient, so we're at 63 percent high, 37 percent moderate.

MR. AMIN: So the typical way that we handle the related and competing conversation is typically after the measure has been endorsed. I think there is a reasonable question, so, Karen, in terms of the use and usability I guess the question sort of is this a use and usability question or this truly sort of a related un-competing measures question that you're raising?

I mean, it obviously could be both and we can welcome response from the developers here on the relationship between the two measures. And I believe CMS is on the line as well since they're at least the co-steward, I believe, of the other PQI measure. So I would welcome that feedback as well. I'm just trying to understand where to center this conversation.

MEMBER JOYNT: It's probably more of a competing measures, unless you think that one is better for usability per se. But I guess I was thinking about it as sort of an implementation, which I guess is different than usability.

DR. DRYE: And I'm just going to disclose that I'm a small part of AHRQ's contract to maintain the quality indicators, which include those measures by Stanford University who leads that effort.

So they overlap. Let me just highlight the differences. You pointed out some of them, Karen. So the outcome -- the ones that

are relevant here, and I'll just -- yes, we'll talk about CHFs. So there is one and it's used in the ACO Program now and for the 2016-2017 measure set. The outcome is just heart failure admission, so it's a narrow outcome. We look at admission for any acute unplanned cause. So as Erica pointed out, we look at three times as many admissions trying to capture things like admissions for pneumonia or other causes that can be reduced in this vulnerable population by better quality of care.

The modeling is really different. It's a single-level model, logistic model just adjusted for age and sex. They were designed as area indicators, so it's just a very -- one thing that the ACO Program did is they narrowed the population. They didn't just take AHRQ's quality indicator. They narrowed the eligible population so the denominator for the measure just to patients with heart failure.

So that measure does overlap with our measure and it's different than our measure. It

gives slightly different information. That was a programmatic decision. In the ACO Program there's also the AHRQ PQI adapted for the program for COPD.

CO-CHAIR BULGER: Any other comments around usability and use we haven't already talked about? Okay. On the phone, Frank?

All right. We'll vote.

MS. HERRING: Voting is now open for usability and use for Measure 2886. Your choices are one high, two moderate, three low, or four insufficient.

Just waiting on one more.

CO-CHAIR BULGER: Still need one more?

MS. HERRING: Yes.

CO-CHAIR BULGER: All right. Everybody want to click again?

MS. SHAHAB: Tom, can you please send me your vote?

MEMBER SMITH: I sent it once. I'll send it again.

MS. O'ROURKE: Tom, we're voting on

usability and use right now.

MEMBER SMITH: Yes, I sent it twice. It looks like the chat room's not working. Do you want me just to tell you or text you? Does it matter?

MS. O'ROURKE: You can tell us.

MEMBER SMITH: It's between zero and two.

CO-CHAIR BULGER: Okay. Thanks.

MS. HERRING: And the results are 5 high, 14 moderate, 1 low, 0 insufficient, so 25 percent high, 70 percent moderate, 5 percent low.

CO-CHAIR BULGER: Any other comments on this one? All right. Let's vote.

MS. HERRING: We're now voting on overall suitability for endorsement for Measure 2886. Your choices are one yes, two no. And we're looking for 20 votes.

MS. SHAHAB: Tom, do you mind just emailing me your vote?

MEMBER SMITH: Yes, this is a yes for me, too.

MS. SHAHAB: Okay. Thank you.

MEMBER SMITH: Is it, Zehra?

MS. SHAHAB: This is Zehra, yes.

MEMBER SMITH: Zehra. All right. I'll try sending you an email. See if you get it.

MS. HERRING: The results are 19 yes, one no. Ninety-five percent yes, five percent no.

MR. AMIN: So before we move on to the next measure, let's handle this related and competing measures issue.

So as I sort of described yesterday, the way that we'll handle this is we will ask for the -- I'll ask the Committee whether or not based on this discussion that we've had whether we believe that this measure is related or competing to another measure in the portfolio. If there's any one particular -- all I'm looking for is one hand to say we should have that conversation.

We need to pull both measures and have

a more robust conversation about both measures and the differences in the specifications. So we will likely have that conversation during our post-comment call and have a discussion about the two measures informed about both measure specifications and have a discussion around the level of harmonization. And if we need to, if they do classify as competing measures, potentially have a discussion around best in class.

So with that, I'll just ask for one hand to say that if we need to have that conversation based on what we've heard today.

MEMBER JOYNT: Yes, I think we need to have that conversation.

MR. AMIN: Okay. So then we'll flag that for a future conversation. So again, just for the record that is for this measure along with the AHRQ PQI measure. Do we have the measure number, quickly? That's all right. So, and it's an AHRQ PQI measure for CHF.

MEMBER BRUCE HALL: Measure No. 0277.

MR. AMIN: Thank you, Bruce.

CO-CHAIR TRAVIS: Okay. Well, thank you all for that great conversation. I do imagine that a lot of the conversation that we just had would also be applicable to the diabetes measure, which is 2887, Risk-Standardized Acute Admission Rates for Patients with Diabetes.

However, we want to be sure that, one, if there are any particular issues related to diabetes that perhaps we did not cover and congestive heart failure, because they're different, let's be sure and focus on that.

And if there is still some lack of clarity or some concern around some of the issues we discussed already, we do want to be sure that we circle back to those. But hopefully we can move through this one a little bit faster since we've had a conversation on the underlying methodology that was for congestive heart failure.

So our discussants for this particular measure are Kathy, John and Keith, but before we

do that I wanted to see if the measure developers wanted to say something particular related to the diabetes measure to kind of ground us in this particular issue.

DR. DRYE: Just a couple quick comments. I don't know if --

Kasia Lipska, are you on the phone? She's our expert endocrinologist who led this measure, but she's traveling to a conference today.

We have a lot more cases in this measure, so that's just a note. There's more of a range in the illness burden of the patients in this measure. As a result, our risk-adjustment does even more to distinguish among them. And then we use a Diabetes Severity Index variable to help us with that. That's validated in claims data. So otherwise, I don't think there's anything fundamentally different about the measure.

MS. SHAHAB: Elizabeth, I did want to let you know that Kasia is on the phone and she

has an open line.

DR. DRYE: Kasia, did you want to add anything? We can't hear you.

DR. LIPSKA: Can you hear me now?

DR. DRYE: Yes. Go ahead.

DR. LIPSKA: So I am driving and it's very busy, so I'm going to have a hard time contributing, but I've been listening.

I think that the things that I wanted to highlight about the diabetes measure are that patients with diabetes are obviously also at high risk of admission. That risk is lower than the risk for admission for patients with heart failure and multiple chronic conditions, as Elizabeth already mentioned.

And one thing you'll notice when we present the measure is that the model that is higher than it is for MCCs and for heart failure. It's about 0.22. And that may have to do with the fact that there is a variation in the population at risk for hospital admissions. There are patients who are healthy patients with

diabetes and then there are those who have a advanced disease and are very vulnerable to hospital admission. And our risk model appears to allow us to discriminate between these populations.

CO-CHAIR TRAVIS: So are we ready to move on, Elizabeth?

DR. DRYE: Yes.

CO-CHAIR TRAVIS: Okay. All right. Well, let's move on to evidence. And I'll go to Kathy first to see if you have anything that you'd like to share with the group. And then we'll go to Keith and then John.

MEMBER AUGER: I think that this measure is very similar to the heart failure measure. They have the same conceptual model of how improving care for chronic conditions such as diabetes might prevent hospital admissions.

CO-CHAIR TRAVIS: Keith, anything to add?

MEMBER LIND: No, I agree.

CO-CHAIR TRAVIS: John, anything to

add?

CO-CHAIR BULGER: No.

CO-CHAIR TRAVIS: Okay. All right. Any questions or discussion from the Committee?

Seeing none, we'll go to vote.

MS. HERRING: Voting is now open for evidence for Measure 2887. Your choices are one yes, two no.

And the results are 19 yes, 0 no, so 100 percent yes.

CO-CHAIR TRAVIS: Okay. Thank you.

Our next criterion is opportunity for improvement or gap.

Keith, anything you'd like to mention?

MEMBER LIND: Yes, there is clearly room for improvement here. And it's similar. I guess the gap is narrower than it was for heart failure. So 45 percent no difference, 40 percent had better than national, and 16 percent worse than national. Definitely room for improvement. So I mean, I would -- that's enough for --

CO-CHAIR TRAVIS: Okay. Thank you,

Keith.

Kathy, anything to add?

MEMBER AUGER: Just note that there are disparities that they talk about as well. In terms of the AHRQ SES Index and as well as the dual-eligible you see differences in performance, which also speaks to gap.

CO-CHAIR TRAVIS: Thank you. John, any additional information?

CO-CHAIR BULGER: No, I would agree with what they just said. I mean, I think with many of the other measures we look at generally there's a lot more people that are in the middle and lot less on the tails, whereas in these measures there were -- the fact that such a -- that I guess it was 40 percent, almost 40 percent were better than the national rate is pretty -- is a lot with what we're used to looking at. But there's clearly a large gap between high and low on this.

DR. DRYE: Yes, I can just add that the number of admissions is just lower also to

begin with, so I think that affects the variation that we're seeing. It's a median. The national crude rate is 41.4 per 100 person-years and in ACOs it's 39.6, so -- compared to heart failure, which is 85 nationally and MCCs at 72. So there's not quite as much of a room for absolute range of number for admissions for person-year.

CO-CHAIR TRAVIS: Thank you for that.

Okay. Any questions or comments from the Committee about performance gap?

Okay. We're ready to vote.

MS. HERRING: Voting is now open for performance gap for Measure 2887. Your choices are one high, two moderate, three low, or four insufficient. We're looking for 20 votes.

The results are 7 high, 13 moderate, 0 low, 0 insufficient, so 35 percent high, 65 percent moderate.

CO-CHAIR TRAVIS: Great. Now we'll move onto to reliability. And, Kathy?

MEMBER AUGER: Sure. So they also did the split-half correlation, and the ICC of that

was 0.89. So there was a high degree of reliability.

CO-CHAIR TRAVIS: Anything to add, Keith?

MEMBER LIND: I would just reiterate my point about looking at other areas, particularly for diabetes outside of inpatient admissions and reiterate Wes' point about mortality. I mean, as we discussed yesterday, if you're not in the sample, it may be that you're dead, which is a significant outcome. That's a bad outcome.

And the fact that there's a parallel measure -- I understand there's a different measure for mortality, but we talked about this yesterday. So sometimes it could skew the way this measure looks. Even though you have another measure, it's difficult to put them side by side for an individual institution and see how they -- is there an ACO mortality measure?

DR. DRYE: No, not exactly. So there are a set of measures around diabetes care, but

patients who die, they're only the -- I mean, our denominator only includes patients that are alive, so it's different in the sense that we use person-years. So we don't count the time after which something had happened to them. But the ACO measure set does not include a mortality measure per se. It includes specific composite around achieving diabetes care. And it's a good point.

MEMBER LIND: Yes, I guess that makes it that much more important. I think the days of acute care that they used yesterday used a -- well, you use a year, so if you use person-years, if you block -- if mortality reduces the number of person-years available -- I don't know, I shouldn't try to figure out how to do it, but --

DR. DRYE: Yes, it's if you die within a year, you -- so if you were only alive through February, you would just contribute 2/12ths of a person-year. You contribute a fraction of a person-year to the denominator. So we don't count that period as being exposed. That's

different. That's consistent with EDAC. It's different than the hospital readmission measures.

But the hospital readmission measures, as you point out, are reported with a mortality measure. We don't have the same thing in the ACO measure set.

CO-CHAIR TRAVIS: John?

CO-CHAIR BULGER: I don't have anything to add to what was already said.

CO-CHAIR TRAVIS: Okay. Any other questions or comments from the Committee on reliability?

All right. Well, I think we're ready to go for a vote.

MS. HERRING: Voting is now open for reliability for Measure 2887. Your choices are one high, two moderate, three low, four insufficient. And we're looking for 19 votes.

The results are 2 high, 17 moderate, 0 low, 0 insufficient, so 11 percent high, 89 percent moderate.

CO-CHAIR TRAVIS: Okay. Now, we'll

move to validity. Keith, any comments?

MEMBER LIND: I mean, they did pretty strong validity testing and used -- I think they used three different methods, right, and scored pretty high on those. I don't know if anybody else had any other comments on this.

CO-CHAIR TRAVIS: Kathy?

MEMBER LIND: They decided not to include the SDS factors.

CO-CHAIR TRAVIS: Thank you, Keith. I'm sorry.

Kathy, anything?

MEMBER AUGER: Yes, overall this model fit better than the previous model with R-squared of about 22 percent exclaimed. I will comment that when you look at the models with and without the sociodemographic factors, it looks like there's a little bit more movement than there was in the heart failure ones, but still for the same conceptual reasons. And because the performance of -- there are some high-performing ACOs even in the high-percentage share of the poor ACOs --

that they still argued to not include them.

I think the one question that I had for the developers is so this, like the other measure, is about planned -- is excluding planned admissions. And I was looking at the planned admission. It looks like debridement of wounds is considered a planned admission, which conceptually for a diabetes measure it seems like wound prevention would be a big thing that you want physicians to focus on. So I was curious as to why that was considered planned and then therefore excluded.

DR. DRYE: Yes, great question. So the algorithm that we used to identify planned admissions we adapted from what we used in the readmission measures, because when we developed them initially, we were thinking about planned admissions broadly, not just readmissions.

And we did make a couple of modifications. We started with the version 3 of the algorithm. So for example, we pulled out -- we don't count as planned amputations. But

around debridement there was -- I don't know the specifics off the top of my head. I can get them for you.

We did a validation of the algorithm. And if you're admitted for an acute condition like sepsis and wound infection is the cause, it's not called planned. If you're admitted for anything acute as a primary diagnosis and the procedure of wound debridement occurs, it will not be called planned. But if you don't have any acute diagnosis and you're admitted for wound debridement, when we went and developed an algorithm that was found to be most often just routine care, which can be good diabetes care.

It's not going to classify these perfectly as planned or unplanned, but the balance was towards those being planned. But we did go through for this measure and just review in this context whether we were striking the right balance. And we made some adjustments. I'm sure it's not perfect.

MEMBER AUGER: I will just kind of

echo that, because I was on a committee where we did chart reviews of these cases. And so, when it wasn't associated with an acute diagnosis like cellulitis or acute osteomyelitis, all of these other wound care procedures were planned follow-up procedures, so staged procedures for wound care that we felt were part of quality care and not necessarily should we be including them in this measure. We shouldn't necessarily be dis-incentivizing that care.

DR. DRYE: That makes sense. Thank you.

CO-CHAIR TRAVIS: John, anything else?

CO-CHAIR BULGER: No. I mean, it still is a little bit of a question in my head when almost 40 percent, I think it's 38.5 percent, end up better than the national average. How valid the measure is just from a face validity standpoint, it just seems like -- and again, I think that's partly because the ACOs you're looking at are a self-selected group and honestly you don't put yourself in a program

unless you think you can do okay. So I think that's part of it, but it's a little bit of a concern.

DR. DRYE: Yes, and I'm definitely learning from this experience that when we put these results in the application, we were very cognizant of how many outliers there are. Like in the more recent data we ran with the AHRQ SES testing update where we had 220 ACOs from the 2013, there's even more outliers. The reason is we have such high -- so if you think of the national average just like a point, a line, it's just the crude rate over all the diabetes patients who are in fee-for-service. What was the number of unplanned admissions?

And so, if that's sitting at 41 and you have very small confidence intervals around our risk-adjusted estimates -- and we have such high volume that we have a nice distribution of risk-adjustment estimates, so we don't have shrinkage or anything like that happening because we have such volumes. It could be that many of

these are just barely over the line in one way or the other, and some are way over the line, so that on a more informative -- or it might -- there is a distribution of the measure scores.

We converted it into a categories because that's how CMS traditionally has reported for consumer interpretation the hospital measure results in this setting. And actually CMS isn't using these categories, and I'm not convinced it would be the right way to report these results because if you're given the statistical power we have, just being like a tiny bit lower than the national crude rate really is probably not meaningful.

So I just want to caveat that that's one way to report the measure score. It has its down sides. I don't know if others --

CO-CHAIR BULGER: Well, and as you said, that's not the way they're using it right now. So you chose to report it that way for us because that's just the traditional way it's been done.

DR. DRYE: Well, and the irony is that a limitation of our models in the hospital settings is we don't identify very many statistically significant outliers because of the approach to modeling and the small sample sizes. So, and CMS likes to use this very high confidence level of 95 percent, which you wouldn't necessarily need to use for reporting the way you might in a research setting where you're rejecting a hypothesis or something.

So we were like, great, we have so many outliers, but as we've been thinking about it more, it maybe conveys more sort of better and worse that we really should be conveying. We just have so much data that we are probably -- it may not be the best way to present the results.

DR. SPATZ: Because of the self-selection bias we anticipated that potentially all 100 percent could be performing better than the national rate and kind of looked at the data a little bit differently at the other end, which is to say, wow, we were surprised to find that

there were ACOs that were worse than the national rate given that they were the earliest of adopters.

CO-CHAIR TRAVIS: Okay. Paula?

MEMBER MINTON-FOLTZ: Sorry if this was covered already, but are we talking about attributed, designated or both for ACOs?

DR. DRYE: I think you're asking how our patients attributed to the ACOs. So we were lucky to be able to use CMS' ACO assignment file, which was post-all of that work, to assign our patients and our data to ACOs, but that file used the policy in place at the time for the Medicare Shared Savings Program and also the Pioneer ACOs.

In the MSSP it's post -- it's attributed based on where they're getting the majority of their care. And for this particular analysis it used criteria that have since been revised a little bit.

CO-CHAIR TRAVIS: Okay. Any other questions or comments about validity?

(No response.)

CO-CHAIR TRAVIS: Okay. I think we're ready for a vote.

MS. HERRING: Voting is now open for validity on Measure 2887. Your choices are one, moderate; two, low; or three, insufficient. We're looking for 20 votes.

(Voting.)

MS. HERRING: The results are 17 moderate, 3 low, 0 insufficient, so 85 percent moderate, 15 percent low.

CO-CHAIR TRAVIS: Okay. Now we'll move to feasibility.

Kathy, any thoughts on feasibility?

MEMBER AUGER; It's very similar to the other.

CO-CHAIR TRAVIS: Thank you. Keith, any additional comments?

(No response.)

CO-CHAIR TRAVIS: John?

CO-CHAIR BULGER: No.

CO-CHAIR TRAVIS: Okay. Any comments or questions from the Committee?

(No response.)

CO-CHAIR TRAVIS: Okay. Ready for a vote.

MS. HERRING: Voting is now open for feasibility for Measure 2887. Your choices are one, high; two, moderate; three, low; or four, insufficient. Again, we're looking for 20 votes.

(Voting.)

MS. HERRING: The results are 15 high, 5 moderate, 0 low, 0 insufficient, so 75 percent high, 25 percent moderate.

CO-CHAIR TRAVIS: Okay. Now we'll go to use and usability.

Keith, any comments?

MEMBER LIND: It's not currently in use, but it's planned for use and seems usable and understandable.

CO-CHAIR TRAVIS: Thank you. Kathy, any additional comments?

(No response.)

CO-CHAIR TRAVIS: None? Okay.

John?

CO-CHAIR BULGER: Yes, I just can't help myself. I think the only thing that's ironic is the comment you just -- and we're talking about use in the -- that you made the comment that the difficulty -- because CMS requires 5/95 for reporting, which is true, and they require 49/50 for payment. I mean, the O/E ratio is the difference between the 49th percentile and the 50th percentile. You become -- it changes your payment on -- not on this measure, but on a bunch of the other measures, whereas for reporting they require 5/95 to say that you're better than the national average or worse than the national -- my soap box. Sorry.

CO-CHAIR TRAVIS: Thank you for your soap box. But it is a recurring theme, so it is part of our thought process.

Okay. Any other comments or questions about use and usability?

(No response.)

CO-CHAIR TRAVIS: Okay. We're ready for a vote.

MS. HERRING: Voting is now open for usability and use for Measure 2887. Your choices are one, high; two, moderate; three, low; four, insufficient. We're looking for 20 votes.

(Voting.)

MS. HERRING: The results are 5 high, 14 moderate, 1 low, 0 insufficient, so 25 percent high, 70 percent moderate, 5 percent low.

CO-CHAIR TRAVIS: Okay. So now we're to the point where we will be voting on overall suitability for endorsement. Are there any final comments or questions from the Committee? Wes?

MEMBER FIELDS: Yes, I've been struggling with this for about the last 10, 15 minutes, and this will probably disturb Taroon because I really feel like our last question should include utility. And so, I'm a little troubled by the utility of this.

I understand the decisions you made. They're defensible from a methodological approach, but I have to believe the most important measure for these very important

populations for an ACO is per-member per month cost and the impact on shared savings.

My bias is that I'm an acute care provider, but I see tons and tons of people, especially in the category of diabetes, and also renal disease, although it's not one of your measures, who come from accountable care entities, and some of them do a much better job than others in terms of managing things.

So for example, for me the difference between a scheduled debridement or a scheduled amputation and an unscheduled one is moot. And I don't think it matters that much to the patient who has their foot cut off either.

And so, I'm just troubled a little bit by what seems like an arbitrary focus on acute care services when what we should all be striving for is the best outcomes for this population in which all costs matters and all interventions matter.

CO-CHAIR TRAVIS: Any comments?

DR. DRYE: No, I mean, I think you're

raising valid points. What we were trying to do was capture an underlying quality domain of care coordination and care efficacy. We weren't really focused here on resource use per se, so we tried to walk that line and we didn't want to generate adverse consequences like discouraging routine care.

So I think that the difference -- I think the supplements, the broader numbers that ACO had on their savings -- and their biggest costs are hospital costs for sure, but here we're trying to get down in all three of these measures to reflecting the quality of care. And that was the design. Is it perfect? No. And those are some of the things that we really struggled with. But it isn't a resource use measure. That would have been done very differently. So it's risk-adjusted. It's pulling out things that we think aren't related to quality. It's not perfect.

MEMBER FIELDS: In a consensus-driven organization like NQF a reasonable compromise would probably be to consider all admissions.

DR. SPATZ: Yes, I will say with the three measures in parallel we had a lot of stakeholder input, and this particularly came up in the multiple chronic conditions, which we'll hear about next. But we had a lot of pushback by some people to say hospitalizations for these very difficult patients are not all bad, even some of the acute ones that we don't classify.

It provides an opportunity to reconfigure resources. Sometimes a break for families, we heard. People were afraid that we might dis-incentivize placing an ICD or appropriately caring for them the amount of duration needed because those things -- and those things are high-quality care.

So we didn't just kind of automatically assume that all -- just don't -- let's just cut the line at acute admissions, but we did hear very loud and clear from a large stakeholder group that we're not looking to reduce all admissions and that some of these are indicators of quality. And that kind of caused

us to -- especially with the diabetes measure, the amputations, because these are outcomes that we really care about, that we really want to avoid.

We want to avoid end-stage renal disease. We want to avoid amputations. But we also want to incentivize good care for people whose diseases have been manifesting for the last 20, 30 years and at this point we're not going to prevent them from going on dialysis, but we can change the way that that's done and improve their outcomes in all.

So we tried to walk that line, but you're right, I mean, there's not like a clear bright line between what should count and what shouldn't count.

MEMBER FIELDS: Well there is in your measure.

DR. SPATZ: There is, right. We had to draw a bright line, but clinically I think we can all come up with cases that don't fit that bill.

CO-CHAIR TRAVIS: Paul?

MEMBER HEIDENREICH: I would just say there's a lot of acute care that's good quality, too. So you're theoretically providing a disincentive to admit someone with chest pain in the emergency room, depending on their -- so I don't think it's a fine line that bad quality is acute and planned is good.

CO-CHAIR TRAVIS: John?

CO-CHAIR BULGER: Yes, I do hear what you're saying, and I think probably you're right, when you talk to -- get TPs together and they talk about they're worried about people withholding care because of this. But to me the balance to that in these measures is, as we have talked about, the end-all be-all for the ACO is total cost of care.

So to me these measures are balancing the ACO withholding all care, because that's the way you could really win in an ACO is not spend any money. And this measure is looking -- so looking at all admissions with this measure would

be the complete balance to what the ACO is after if they were in a perverse environment, which would be not providing care. So looking at all admissions. I don't think looking at all admissions in this group is going to make people not admit someone for a reason, that they have a reason not to admit them, which is because it's going to be a cost to their ACO.

CO-CHAIR TRAVIS: Karen?

MEMBER JOYNT: Yes, I have to respectfully disagree that I don't think this measure balances the risk of not doing anything. If anything, it makes it worse because it just puts more -- puts not only on your cost side, but also on your quality side not to do anything. In theory the ACO measures like diabetes control and other things where you're measuring LDL, those should incent you to do more to make people better, but there isn't -- to the point raised earlier, there isn't an offsetting -- like keeping people alive and out of the hospital in a way that is in keeping with their goals and not

in doing too little that exists currently. Maybe that's a place for future measure development is some sort of offset to some of the risks we worry about, thinking back to capitation and things like that.

CO-CHAIR TRAVIS: Helen?

DR. BURSTIN: I was actually going to raise the point about balancing measures as a concept we've talked a lot about as you look at some of these more population-oriented measures. And in fact, this also relates to a potential competing measure discussion. But there are also AHRQ Prevention Quality Indicators, specifically in the diabetes space, that get at long-term complications admission rate, short-term complications admission rate, the lower extremity amputation rate among patients with diabetes, as well as the uncontrolled diabetes admission rate.

So in some ways those -- but they're not part of this program, which is a separate issue, but I think it is a question as we come back and look at competing measures. If some of

those measures are out there, one consideration would be how do you then factor in concerns about measures that they push the bubble in one direction about having some way of understanding if there are balancing measures that give the full picture? And obviously something much more aligned with what we tend to do at the MAP tables, but I think this is obviously -- raises a potential concern for this measure.

CO-CHAIR TRAVIS: And I'll just comment on that as well, because as you were talking of course the whole MAP conversation came to my mind about looking at the programmatic portfolio and having this rich discussion about what are the drivers in an ACO? And the financial driver is critical and is key to that.

So I think it's very good conversation to be sure the MAP understands that our concerns were around do we have some measures in the portfolio for the ACO Program that are balancing measures to be sure that care is not being withheld. Because the financial incentive is to,

quote/unquote, "not spend the money," which could mean not provide the care.

So I do think that it causes us to have to look at this program a little -- I mean, an ACO-type measure; not program, but ACO measure, a little bit differently than we might in a fee-for-service-type environment where the financial incentive traditionally has been to do more.

So I think let's find a way; and Erin and I were talking about this earlier today, to be sure that this kind of a conversation we're having gets not just on our side, on the endorsement side, but migrates into any discussion that goes on at the MAP side. Because I think this has been an important aspect of what we're evaluating and struggling with, quite frankly, relative to these measures.

Leslie?

MEMBER LESLIE HALL: I just wanted to comment. I don't think we want to end up with whack-a-mole for measures, and it feels like that

when you're trying to juxtapose all of these things. And it's worth discussion further.

CO-CHAIR TRAVIS: No, I agree. I mean, we looked at the portfolio earlier that's in this Committee's, and we went through it rather rapidly. There may be some time that it makes sense for us to have not an in-depth, but at least a little bit -- maybe a different way of picturing it so that we can understand what kind of measures are in here. Because I think you bring up a really good point, I mean, just continuously having more and thinking about how they're related. And I think it's an issue that we're beginning to see.

So thank you, Leslie, for bringing that up.

Yes?

DR. DRYE: I think this is really interesting and I think our communities' view is evolving on this somewhat. So I would encourage you to think about it across actually the readmission -- this is the Readmissions

Admissions Committee, but across the readmission measures as well. It was really in the context of the readmission measures that we had a lot of discussion at NQF and elsewhere around how to evaluate admissions and should we be looking at only related admissions -- and I'm sure you guys are doing all this -- only related admissions that are directly related? Now, you had your hip/knee replacement. Did you come back with an infection or a DVT? Or should we look more broadly?

And we really evolved as a community towards taking a very patient-centered view saying we want to look at all admissions that could be related and then risk -- we don't think it's going to be zero, and we risk-adjust for differences that will take care of the unrelated things like you walked across the street and got hit by a car. Hopefully that wasn't because you were over-medicated. We don't know if it's related or not, but we try to get a lot of that noise out of the way and hold onto that broad

set. But we didn't go all the way to counting everything because there was a strong view among many stakeholders that many admissions aren't providing a quality signal.

So we evolve. We approach these admission measures starting really where the debate was at the time, which is we're trying to capture admissions related to quality as broadly viewed in the setting of primary care, chronic disease management from the patient's perspective. And an admission for a hip replacement in that context, if you happen to be a patient with a diagnosis of diabetes, is not a quality signal.

So it's those concerns and that context and those discussions over a number of measures, not just these that kind of led us away from saying, well, an admission is good or bad. We're trying to craft a measure to reflect quality of care, not to count admissions. And I'm hearing maybe we haven't struck the right balance. There are other measures that do it

differently that are essentially focused on more narrow highly-related things for narrower cohorts depictions.

So it's a great discussion. I mean, I'm agreeing with you, but I think it's broad beyond these measures.

CO-CHAIR TRAVIS: Yes, Carol?

MEMBER RAPHAEL: In line with that; and this is just for the future, I think one of the issues I can relate to in home healthcare is you get a discharge for someone who's had a hip replacement and you in and the person has out-of-control hypertension, they have diabetes that's not well managed, they have extraordinary pain from arthritis and they're depressed. And you go back to the physician, the orthopedic surgeon who says it's not my job. You've got to go elsewhere to deal with all the rest of that.

And what you're trying to do is get the best outcomes for someone who has multiple chronic conditions at the same time that you're trying to prevent the readmission for some

infection or decline having to do with the hip replacement. And you're working in two domains. The first one is measured; the second is not measured, and it's much more complicated. But to me that is part of how you have to think about people who have multiple chronic conditions.

CO-CHAIR TRAVIS: Thank you, Carol.

Wes?

MEMBER FIELDS: Yes, I want to go back to something Karen raised earlier. And for me I bow down to primary care-sensitive conditions, but I think the reality is that for things like diabetes or renal failure -- you talked yourself about sort of inheriting the healthy 65-year-old with a fasting blood sugar of 130, where you actually have an opportunity to prevent comorbidity over a period of time.

But the reality is that the people at the other end of the rainbow treating a very advanced disease is the likelihood that they're even able to get through the door of a standard lower-tier medical home is low. And they in fact

have a very different set of providers. Hopefully the ACO is enlightened enough that they're well-integrated, but that's not always the case. And a lot of what you're struggling to measure is when it's not.

But I think the reality is that the more complex or advanced these diseases are, especially the MCC patient, the more likely they're in a different orbit between acute care facilities and post-acute care. And that's one of the reasons I think it's a mistake not to measure all bed days, because they all matter. They matter to the patient. They matter to CMS. They matter to providers.

So I think some of the distinctions you made about whether care was scheduled was not ultimately are arbitrary and don't reflect anything that's worth measuring.

CO-CHAIR TRAVIS: Okay. Any other comments? Anybody including the developers want to say anything before we finish?

(No response.)

CO-CHAIR TRAVIS: Okay. So we're ready for the vote.

MS. HERRING: Voting is now open for overall suitability for endorsement for Measure 2887. Your choices are one, yes; two, no. And we're looking for 20 votes.

(Voting.)

MS. HERRING: The results are 18 yes, 2 now, so 90 percent yes, 10 percent no. Thank you.

CO-CHAIR TRAVIS: All right. Well, thank you. I'm going to turn it back over to John.

MR. AMIN: Before we move onto the next measure, Cristie, if it's okay, based on our prior conversation it appears that we'll probably need to have also a conversation around related and competing measures for this measure as it relates to 0272, the Diabetes Short-Term Complications Admission Rate, the PQI 01, along with 0724, the Diabetes Long-Term Complications Admission Rate, PQI 03, and then 0638,

Uncontrolled Diabetes Admission Rate, PQI 14.

Is there anyone -- it sounds like there's agreements about that, so we will flag that for follow-up for conversation with the Committee.

And so, I'll turn it back over to John.

CO-CHAIR BULGER: Okay. Thank you. And I did want to say to -- I got discombobulated in the point I was trying to make there. And to Karen and Helen's point and to somewhat what was just said, I think the universe of ACOs is very heterogeneous in the way it sits right now with the shared savings and essentially comparing yourself to yourself, at least in the old method, now realizing that the new rule that came out in the last couple of days changes that a little bit. It still offers the availability for people to -- and I hate to use the word "game," but you can go into a system where there's a very large Medicare cost beneficiary, kind of a McAllen place, and there's money to be made essentially.

So coming up with measures -- and I agree completely that these measures don't necessarily do it, but coming up with measures, as Helen talked about, that are able to check that to some extent I think is extremely important for us to do.

So the next measure is 2888, Risk-Standardized Acute Admission Rates for Patients with Multiple Chronic Conditions. And again, it is much like the two we've just talked about, but somewhat different. Frank, Mae and Steve are the discussants. And we'll start with the developer.

DR. DRYE: Okay. I'll try to cover quickly. You guys have given us thorough discussion.

So this was a fun to measure to work on. It was a privilege to work on it because as you know there are very few measures for patients with multiple chronic conditions. So I just wanted to highlight how we put together the cohort for this measure, otherwise, its properties aren't that different from the two

other measures.

We started by looking at frameworks on what do we mean by patients with multiple chronic conditions including NQF's, and NQF's MCC framework defines MCC patients as having two or more concurrent chronic conditions that act together to significantly increase the complexity of management that affect functional roles, health outcomes, compromise life expectancy or hinder self-management.

So we started there. Sought expert input including from Mary Tinetti at Yale and Cynthia Boyd at Hopkins and tried to really -- looked at the frameworks that are coming out of the geriatrics expertise and defined a list of chronic conditions consistent with those definitions and then empirically looked at how they were acting together to contribute to risk.

And we settled on -- we developed an approach and gave two versions of it put out in public comment using eight fairly broadly defined groups of chronic conditions. We said we could

define our cohort using two or more of these or a more restrictive narrow cohort would require three or more of these conditions, which was consistent with at least one of the clinically-driven frameworks.

And it was just really interesting to get expert input and public comment. And in the big picture we ended up defining this cohort as patients with two or more of eight chronic conditions. They're listed -- or I can list them for you quickly. Ischemic heart disease or history of MI, Alzheimer's disease and related disorders of senile dementia, chronic kidney disease, chronic obstructive pulmonary disease, depression, atrial fibrillation, diabetes, heart failure and stroke. Sorry, we pulled out -- we didn't use diabetes. I mentioned that before. And we also looked at non-traumatic joint disorders or arthritis, which is an important chronic condition, but it doesn't contribute much in our analyses to admission risk.

So when we put this out in public

comment, some of what we got back was just acknowledgment from primary care providers and internists that they felt like our two-plus or two or more chronic conditions cutoff was as good one. That represented about 25 percent of Medicare fee-for-service patients with any chronic condition. And that just kind of -- it was consistent with their clinical experience, that they felt about a quarter of their patients with chronic conditions really fit these framework definitions of multiple chronic conditions. That left us with a pretty big cohort nationally of 4.9 million patients. And within ACO it was about 240,000 patients.

There's no one way to do this, but I think we've drawn a line systematically with expert input. And I'm happy to answer questions about how we came to this particular group of patients.

CO-CHAIR BULGER: Okay. Thank you. So, Frank, you want to --

MEMBER BRIGGS: Yes. So as mentioned,

all the previous discussion I think is relevant. These are patients, again unplanned admissions with two of those chronic conditions listed representing a number of Medicare fee-for-service populations.

The strategies for improvement are basically the same for interaction with the ACOs: building better social networks, better chronic care management and those things. What the developers have started to touch upon was that list of eight. Unsure really at this point how the eight came about as well as the interaction between the eight. So do the eight all interact equally? So does someone with AMI carry the same risk as AMI, Alzheimer's, Alzheimer's, AFib, that type of interaction between the two? But other than that I think that the measure itself is very similar to the previously discussed.

CO-CHAIR BULGER: Okay. Mae?

MEMBER CENTENO: Nothing to add. Thank you.

CO-CHAIR BULGER: And, Steve?

(No response.)

MS. O'ROURKE: He's not here.

CO-CHAIR BULGER: He's not here?

MS. O'ROURKE: No Steve.

CO-CHAIR BULGER: Helen, go ahead.

MEMBER CHEN: Hi. It's just more of an editorial comment. I just wanted to thank you for including these conditions in your model. As a geriatrician we struggle with all the time in terms of the concept of multi-morbidity. And these eight conditions are actually pretty strongly associated with this bigger concept of frailty, and I really wish -- actually that's what I would like to see, is a measure for determining how we can help people to manage the concept of frailty in terms of readmissions, because I don't know that there's any really good way of getting at that from claims data, but that's really what we're mostly interested in in terms of this population that is so at high risk for readmission, or just admissions to begin with.

I do think that it is striking that most of these conditions do risk-adjust from -- in terms -- from the CMS HCC kind of risk-adjustment, except for I believe dementia is no longer on the list. So just an editorial comment.

DR. DRYE: Thank you.

CO-CHAIR BULGER: Any other comments?

(No response.)

CO-CHAIR BULGER: Okay. So we'll vote on evidence.

MS. HERRING: Voting is now open for evidence for Measure No. 2888. Your choices are one, yes; two, no.

(Voting.)

MS. HERRING: The results are 20 yes, 0 no, so 100 percent yes.

CO-CHAIR BULGER: Okay. Thank you. And just a quick -- we're close on time. We're not horrible. So just try to keep comments on this measure. And if we don't have anything new from the previous two -- the other thing is we

had a break built in. I think we're not going to take the break, but people have been getting up and taking their own break. Encourage you to do that. Sitting is the new smoking, so --

(Laughter.)

CO-CHAIR BULGER: -- get up and move around, if you get a chance.

So next we have gap, and I'll start with Mae.

MEMBER CENTENO: Yes, so for this measure the gap that was -- so the average, the national risk-standardized acute admission rate for fee-for-service is 71.9. For the ACO it's about 69.3. And the range is anywhere from 62 to 76, so there's quite a gap with this one.

Forty-seven ACOs or forty-one point two percent were better than national rate. And then there is a subset of 45 ACOs that are no different than national, but about 19.3 percent have lower than national. So there's room for improvement.

CO-CHAIR BULGER: Okay. Frank, any --

MEMBER BRIGGS: Just to point out that they did look at SES adjustments similar to the other ones. They found little difference when you risk-adjust with or without the adjustment. And the recommendation was not to adjust for the modifiable factors.

CO-CHAIR BULGER: Okay. Any other comments from the Committee?

(No response.)

CO-CHAIR BULGER: All right. We'll vote on the gap.

MS. HERRING: Voting is now open for performance gap for Measure 2888. Your choices are one, high; two, moderate; three, low; four, insufficient.

(Voting.)

MS. HERRING: The results are 4 high, 16 moderate, 0 low, 0 insufficient, so 20 percent high, 80 percent moderate.

CO-CHAIR BULGER: Okay. Thank you.

Reliability. Frank?

MEMBER BRIGGS: So this is going to be

driven from the administrative claims similar to the others. They did do reliability testing. They used the MEDPAR data as well as the AHRQ SES Index. They did tests and retests with the split-half methodology. I want to say the intraclass correlation coefficient there came out to be a 0.84, which suggests that it was high. I think that's it.

CO-CHAIR TRAVIS: Okay. Mae?

MEMBER CENTENO: Nothing to add.

CO-CHAIR BULGER: Okay. Anything from the Committee? Paul?

MEMBER HEIDENREICH: Oh, wait. Sorry. I want to talk about validity.

CO-CHAIR BULGER: All right. Anything on reliability?

(No response.)

CO-CHAIR BULGER: Okay.

MS. HERRING: Voting is now open for reliability for Measure 2888. Your choices are one, high; two, moderate; three, low; four, insufficient.

(Voting.)

MS. HERRING: The results are 3 high, 17 moderate, 0 low, 0 insufficient, so 15 percent high, 85 percent moderate.

CO-CHAIR BULGER: All right. So we're at validity. Mae?

MEMBER CENTENO: So the developers performed face validity and also used the risk model diagnostics with an R-square of 0.123, so really not as great as diabetes, but close to what they found in heart failure. They also did conceptual modeling of the SES and for reasons already mentioned earlier decided to exclude those.

CO-CHAIR BULGER: All right. Frank?

MEMBER BRIGGS: Nothing to add.

CO-CHAIR BULGER: All right. Anything from the Committee? Larry?

MEMBER GLANCE: This is just kind of a general comment; and we touched upon this a little bit earlier, but we're going to be making the transition from ICD-9 to ICD-10 codes. It's

a massive transition. We're not going to have any idea about data reliability at all for several years, and really a huge number of models, most of the models in fact in the portfolio models of measures that we have we're really not going to know anything about their model performance.

In general, and maybe this is more for NQF, what is the plan? Are we going to go back and reevaluate every single model based on really a very new methodology?

DR. BURSTIN: Yes, so I'm happy to quickly respond to that. We don't have the person in the room who does this, but my understanding is our current approach has been that anybody who has claims-based measures; actually Yale probably knows this better than anyone, has a snip of ICD-9 and ICD-10 during this transition. And again, you're absolutely right, we don't know how this is all going to play out, but at least the thinking is to prospectively ensure that that information is

available as that transition begins to happen.

MS. MUNTHALI: In addition to that they have to submit the plan for converting over to ICD-10, and it's by 2019.

DR. DRYE: Yes, I would just add I think it's a great issue to think about with NQF actually, because what we're doing now is for CMS, with CMS we're thinking about how to do the most robust testing possible because some of these measures are already in public reporting, including these.

So we will look at the consistency of the risk variables, the relationships of those in terms of their frequency, the relationship of those variables to the outcome, what do they look like in -- do they have the same binary relationship? What does the overall model look like? But that isn't really even enough. We're right now still developing our testing plan to really be able to confirm the models that are working in that ICD-10 data.

And I think it's a great issue for --

I don't know that it -- I don't know what context you guys would be involved, but it's a major focus of our work right now.

CO-CHAIR BULGER: Paul?

MEMBER HEIDENREICH: For heart failure and diabetes there was that nice analysis of the SES by quartile and for different -- and how they were labeled as outliers, but I didn't see it for this. Were the results similar or was -- any notable differences?

DR. DRYE: Yes, we put the updated analysis in the memo that we sent last week, and there's some analysis in the original report. It's pretty similar in that in the -- when you use the AHRQ SES Index and nine-digit ZIP code, you have in the quartile of ACOs with the -- of the 220 ACOs, so 55 ACOs with the most low-SES patients, 33 percent are statistically better than the national rate, 31 percent are worse. And that distribution is skewed. There are more who fall into the worst category than in that quartile with the fewest low-SES patients where

11 percent are worse.

So you do see -- I'm going to hand this down to you so you can look at it because it's hard to just -- at least for me, to listen to these quartile results without looking at them. But you do see the same in that there are a lot of -- the 25 percent in the -- I mean, sorry, as I said, 32 percent of the ACOs with the most low-SES patients do very, very well.

So that again, we're advocating for leaving that visible and having that be the benchmark and having those ACOs, which in ACOs is a little more collaborative than in the hospital setting, really being the drivers of change and innovation for those ones that are still struggling. But I'm going to hand this down to you so you can see. And let me know if you have questions.

CO-CHAIR BULGER: Okay. Any other comments on -- Leslie?

MEMBER LESLIE HALL: So building back on the ICD-10 question, when you're doing that

analysis doesn't the whole nature of whether something is planned or unplanned -- isn't that impacted by ICD-10 versus ICD-9, so just that are just that basic? And so, how do you then reconcile that? It seems quite difficult.

DR. DRYE: I mean, I think it would be -- again, I'm looking at Helen because we have formal testing plans. I know other CMS measure developers do as well. And we haven't really publically shared them. I wouldn't mind sharing the outline, but we go through a whole series of steps to make sure that the -- first we code everything in ICD-10, meaning that we look at -- we use GEM's crosswalk, we get expert input, we make sure that we have the right ICD-10 codes that we think people are going to code.

But I think to Larry's point, it's not clear that people are going to use ICD-10 codes exactly how we expect them to. Sometimes a single code is replaced by a double code. And so we're going to be not assuming that those same codes will be used as -- those new codes will be

used as expected or per coding guidance. We're going to be looking to see are we seeing the same patterns of comorbidities? Are we seeing correspondence in the way these risk variables behave?

Beyond that, I mean, I think we would be happy to share strategies on that, how we're going to be doing this testing going forward, but it's kind of a longer conversation.

What would you suggest, Helen?

DR. BURSTIN: Yes, I think it's a longer conversation. We do regularly convene all the measure developers, and it might be a good topic for an upcoming webinar of all the developers to get a sense of where they are in their plans and maybe share and learn what each of them is doing. I know this has been a big issue for AHRQ certainly on all the AHRQ PSIs, PQIs, etcetera.

CO-CHAIR BULGER: o Any other comments?

(No response.)

CO-CHAIR BULGER:

MS. WATT: All right. So we're going to vote on validity. Remember this one the highest you can do is moderate.

MS. HERRING: Voting is now open for validity for Measure 2888. Your choices are one, moderate; two, low; three, insufficient.

(Voting.)

MS. HERRING: The results are 16 moderate, 4 low, 0 insufficient, so 80 percent moderate, 20 percent low.

CO-CHAIR BULGER: Okay. So feasibility. We'll start with Frank.

MEMBER BRIGGS: Yes, so similar to all the other measures we discussed this data will be coming primarily from electronic claims, so your admission, discharge and information from the billing and your ICD-9 coding. So apart the ICD-10 conversion it's all currently there and spelled out.

CO-CHAIR BULGER: Okay. Mae?

MEMBER CENTENO: Nothing to add.

Thank you.

CO-CHAIR BULGER: All right. Any comments from the group?

(No response.)

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

MS. HERRING: Voting is now open for feasibility for Measure 2888. Your choices are one, high; two, moderate; three, low; four, insufficient.

(Voting.)

MS. HERRING: The results are 12 high, 8 moderate, 0 low, 0 insufficient, so 60 percent high, 40 percent moderate.

CO-CHAIR BULGER: All right. And we're on use. Mae?

MEMBER CENTENO: Oh, on use, as already mentioned, this is already used in some of the Medicare Shared Savings with potential use for pay-for-performance beginning 2017.

CO-CHAIR BULGER: Frank, any comments?

MEMBER BRIGGS: Nothing to add.

CO-CHAIR BULGER: All right. Anything

from the Committee on use?

(No response.)

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

MS. HERRING: Voting is now open on usability and use for Measure 2888. Your choices are one, high; two, moderate; three, low; four, insufficient.

(Voting.)

MS. HERRING: The results are 4 high, 16 moderate, 0 low, 0 insufficient, so 20 percent high, 80 percent moderate.

CO-CHAIR BULGER: Okay. And last any further comments from anybody on the overall suitability?

(No response.)

CO-CHAIR BULGER: All right. Seeing none, we'll go to vote.

MS. HERRING: Voting is now open on overall suitability endorsement for Measure 2888. Your choices are one, yes; two, no.

(Voting.)

MS. HERRING: The results are 20 yes,

0 no, so 100 percent yes. Thank you.

CO-CHAIR BULGER: Great. Thank you. I want to thank the developers for sitting through that marathon and helping us with that and for the measure.

DR. DRYE: Thank you.

CO-CHAIR TRAVIS: And I'll turn it over to Cristie.

CO-CHAIR TRAVIS: Yes, and my thanks as well. And it's kind of nice, we're going to actually change topic area and think about psychiatric care. We have a large variety of measures that we consider in this standing committee. So we're going to be looking at Measure 2860, which is 30-Day All-Cause Unplanned Readmission Following Psychiatric Hospitalization in an Psychiatric Hospital.

And we will start off -- once we switch out our developers, we'll start off with some comments from the developers. And our discussants are Frank, Keith -- and is Tom on the phone today?

PARTICIPANT: He is.

CO-CHAIR TRAVIS: Okay. Great.

MEMBER SMITH: You bet I am. I've been waiting three years for a psychiatry measure.

(Laughter.)

CO-CHAIR TRAVIS: Well, I'm sorry you had to wait until the second day for this measure, but I'm glad you're there, Tom.

So we will start with the developers making some opening comments.

MR. CAMPBELL: Sure. Hello, and I'm happy to hear about the enthusiasm for behavioral health. My name is Kyle Campbell. I'm Vice-President of Pharmacy and Quality Measurement at the Health Services Advisory Group, and I'm joined by my colleague Dr. Almut Winterstein from the University of Florida, and we've led the development of this measure for CMS. We've collaborated for the past several years on quality measures for national CMS reporting programs and have a number of measures in the

portfolio that are endorsed.

So for your consideration today is a 30-day all-cause unplanned readmission measure. And this measure was developed for use in the Inpatient Psychiatric Facility Quality Reporting Program, which is a pay-for-reporting program, not a pay-for-performance program. And the facility-level measure estimates an unplanned 30-day risk-standardized readmission rate for the adult Medicare fee-for-service population. In this case this measure does include patients 18 years of age and older because a large number of those patients that are in our population are eligible for Medicare for the other reasons that were talked about before.

The measure includes patients with a principal discharge diagnosis of psychiatric disorder or dementia or Alzheimer's disease and a few key points about the importance of the measure that I want to bring forward.

First of all, readmissions following psychiatric hospitalization have been identified

as a key gap by national stakeholders, including the MAP, and this measure has been conditionally supported by the MAP for the IPFQR.

The second thing is that data suggests ample room for improvement with this measure. So the measure rates for the IPF setting are higher than what we see in the acute care setting with risk-standardized readmission rates of 21 percent versus 15 percent. And we also see more variation in performance, which we would expect given the fact that in the acute care setting we've been working on readmission reduction for quite a while, but in the behavioral setting we have not.

As you know, readmissions are costly and an undesirable outcome for patients and caregivers and we identified ample evidence with regard to the fact that providers could influence measure rates through the adoption of interventions. To that end, we anticipate that a key benefit of measure implementation will be innovation and interventions to address

integration of behavioral and physical care.

This is particularly important for this population which has very early mortality compared to the general population principally from treatable chronic disease and infectious disease. And in one-on-one patient caregiver interviews which we've conducted throughout our measure development process and focus groups we heard loud and clear that care coordination is one of the key gaps in this particular population, and the handoff between the discharge and the outpatient care is extremely important.

Finally, the majority of public comments received from stakeholders support this measure and felt it addressed an important topic. In terms of the specifications, they are harmonized to the extent possible with the CMS readmission measures used in other settings and those that were presented to you from Yale-CORE.

Our expert panel carefully considered appropriate inclusion and exclusion criteria that were clinically and empirically determined. And

like the other readmission measures this uses the planned readmission algorithm to excluded planned readmissions from the outcome.

In terms of the measure testing, I just wanted to let you know that it was very comprehensive. So we obtained the entire National Administrative Claims data set from 2011 to 2014 for measure testing, and that measure score reliability was tested in both split sample and bootstrapping with intraclass correlations found to be moderate to substantial.

For our validity we established the use of diagnosis and procedure codes that are used for billing and that have been validated by other CMS readmission measures. We did a comprehensive risk factor assessment and modeled development including specifically defining psychiatric comorbidity diagnosis grouping based on clinical input and empirical analyses. And we also conducted empirical model validation and a systematic evaluation of face validity by a multi-disciplinary TEP.

The risk-adjustment model includes age, gender, principal discharge diagnosis, comorbidities present during the admission or the 12 months prior and psychiatric-specific risk factors.

Analysis of SDS risk factors was also very comprehensive. We looked at 22 different variables. We did a lot of work to geocode-specific data sets to our data, including at the nine-digit ZIP code level. And unfortunately, as you see these variables did not improve model discrimination. For that reason and several others that I assume we'll discuss, these variables were not included in the final risk model.

The final model has a c-statistic of 0.66 indicating adequate discrimination, which is comparable to risk-adjusted readmission measures used in the other CMS programs.

So in summary, we believe the measure represents a really important outcome for patients with psychiatric disorders and this will

help to support innovation and care coordination in the integration of behavioral health for a very vulnerable patient population. And we'd greatly appreciate your consideration of this measure and we look forward to your comments and questions. Thank you.

CO-CHAIR TRAVIS: Thank you.

Okay. We'll start with looking at evidence. And, Frank, do you have any comments regarding evidence?

MEMBER BRIGGS: Yes, but first I need to at least disclose that I previously worked with both Kyle and Almut on a technical expert panel for CMS around adverse events in the hospital, completely unrelated to this measure topic. I wanted to at least point that out so that everybody on the Committee level was aware.

In terms of evidence, as described, this measure is looking at unplanned all-cause readmissions to the inpatient facility. I did have one question regarding the numerator statement which is described as unplanned

readmission to an inpatient or a short-stay acute care hospital following discharge. The developer further distinguished that. They are looking at days between days 3 and 30 really to I think account for those transfer stations between hospitals and interrupted stays. But my question is really around the short stay being included in the numerator statement, but it is not in the denominator statement.

In terms of evidence and support, I think very similar to other measures that we've discussed today a lot of the same type of interventions can be put in place around intensive care management and connecting patients to services out into the community.

CO-CHAIR TRAVIS: Thank you.

MEMBER BRIGGS: Thank you.

CO-CHAIR TRAVIS: Thanks, Frank.

Keith, anything to add?

MEMBER LIND: Nothing to add. I think they covered it.

CO-CHAIR TRAVIS: Thank you.

And, Tom, anything to add?

MEMBER SMITH: Yes. No, I agree with the developer's comments. I think I'm the lone psychiatrist on the panel here, so I really have been waiting for a psychiatry measure for a long time, and I'm glad to see it.

And the developers are right, we just -- the field has not turned its attention to admissions and readmissions and care coordination issues in behavioral health to the extent it has in general med/surg. So I think this measure is sorely needed, and I'm very happy to see it. The opportunity is huge. We don't have much data on readmissions, but we are seeing now is that they're high and perhaps higher than in general med/surg.

In connections with aftercare, rates of connection with aftercare following discharge from inpatient mental health are very low. In the Medicaid Program it's upwards of -- close to 50 percent of the people don't connect at all in the first month following discharge with

aftercare. So we need to focus more on care coordination. The opportunity is there. And there is a lot of evidence suggesting that care coordination interventions and discharge planning practices can significantly impact on aftercare connectivity and readmission rates.

CO-CHAIR TRAVIS: Thank you, Tom.

I would like to see if the developers mind to answer Frank's question since he raised it a moment ago relative to the inclusion of short-stay acute care.

DR. WINTERSTEIN: Yes, this was actually a double-whammy question, because there's two components that Frank brought up. One is to incorporate acute care hospitals in the numerator and then whether they should be incorporated into the denominator. So I start with the latter.

This measure was specifically focused on the IPF environment. If we had included patients who were the index admissions where the index admission was in the acute care hospital,

that would probably go into a different measure framework within CMS; for example, into IQR. And our charge was specifically to develop something for the IPF, so that's the major reason. But we did comment in our reports that there certainly is a need for patients who would be admitted with a principal diagnosis of mental health disorders in the acute care environment, and it would certainly make sense to expand this measure into this environment, which of course we did not test.

With respect to the numerator statement we did include readmissions into the acute care environment for two reasons: One is of course there may be admissions with a principal mental disorder because patients' comorbidities are too complex to allow readmission into an IPF.

The second is though that we specifically focused on an all-cause readmission measure. So what that means is that we are also looking at patients who might be readmitted for

non-mental principal diagnosis, and these would have of course pretty much exclusively occurred in the acute care environment.

CO-CHAIR TRAVIS: Okay. Thank you very much.

Pam?

MEMBER ROBERTS: So I just had a couple questions. One, so does this include the 5150 people that are forced readmissions or the involuntary readmissions? So, because there could be an access issue when there's only so many in the state. And so then they could be penalized for taking these people and it becomes an access issue because there's nowhere else for these people to go. So was 5150-type patients included?

DR. WINTERSTEIN: Yes, they are included. The claims are not particularly specific with respect to the legal status of patients where the patients were -- there is an admission source variable that looks what whether somebody was admitted from court or from prison,

but the TEP as well as our work group felt -- or found out that this is really not reliable enough to make that distinction.

MEMBER ROBERTS: So then what happens with the facilities that do take these patients? Like for example, in California there's only 20 psychiatric places that will even -- throughout the state. So these people are taking and giving access to these people, but they could be penalized because they have these patients there. I know there's no way on the claims to notify that, but it puts -- I mean, because we have an access issue. Especially in the State of California it becomes a huge issue. And now if we're not going to be able to -- people are going to be concerned about readmissions, I get concerned.

CO-CHAIR TRAVIS: Thank you, Pam.

Leslie?

MEMBER LESLIE HALL: In your conditions do you include substance abuse? Is that part of the primary conditions?

DR. WINTERSTEIN:

(No response.)

MEMBER LESLIE HALL: And so back to Pam's question of access and just what's available 24 hours versus not. So you have many patients presenting into the emergency room, stabilized and transmitted at hours of operation. How does that get impacted or accounted for in this?

DR. WINTERSTEIN: Principal diagnosis as the major reason for admission in the IPF environment was actually surprisingly low. Comorbidity extremely high, but the principal diagnosis is fairly equally shared between psychoses, major depression, bipolar disorder and dementia, Alzheimer's disease.

CO-CHAIR TRAVIS: Thank you.

Derek?

MEMBER ROBINSON: Thank you. I was just seeking some clarity. So when you state in IPF, that does not include a locked unit in a short acute care hospital? That would be purely

a stand-alone inpatient psychiatric hospital?

MR. CAMPBELL: No, there is actually both. So there's stand-alone units, of which I think there's about 50 in our data set, and then there are 1,700 total. And the remaining are units within acute care facilities.

MEMBER ROBINSON: Okay. All right. So what you have in your numerator is referenced to like a medical admission?

MR. CAMPBELL: Correct. Yes, as Almut mentioned, it's all-cause, so --

MEMBER ROBINSON: Okay.

MR. CAMPBELL: But in the denominator it's strictly related to IPF. And IPF is a payment designation, so those units within acute care facilities bill under IPF.

MEMBER ROBINSON: Okay. Great. Thank you.

CO-CHAIR TRAVIS: Paul?

MEMBER HEIDENREICH: Yes, I know theoretically we only need a rationale, which I have to say I think is at too low of a bar for

NQF for these things, but even though we only need a rationale, in terms of actually randomized trials, it seems the best data is for intensive case management, as was summarized by the Cochrane Review. But how much control do the facilities have over starting and implementing intensive case management, which it seems like that would be more of a system that included the outpatients?

MR. CAMPBELL: Yes, that's a good question. Karen, are you available to answer?

(No response.)

MR. CAMPBELL: Perhaps on mute?

DR. PACE: Yes, this is Karen. Can you hear me?

CO-CHAIR TRAVIS: Yes.

MR. CAMPBELL: Yes.

DR. PACE: Okay. So the idea is not so much that the IPF would necessarily be initiating the intensive case management on an outpatient basis, but that would be one type of aftercare to connect patients to if they exist.

So I think the observation about what's available for these psychiatric patients afterwards is a good one, and the incentive of this measure, as most of the hospital readmission measures, is to facilitate the working between the inpatient facility and outpatient services and trying to connect patients to those if they exist.

But you're right that there's no extensive research specifically targeted to IPF facilities, but a lot of the concepts apply equally well to care coordination and facilitating appropriate follow up and services after discharge.

CO-CHAIR TRAVIS: Thank you.

MS. SHAHAB: Cristie, Tom Smith did you want to make a comment as well.

CO-CHAIR BULGER: Hey, Tom?

MEMBER SMITH: Yes, a couple of comments. I think many hospitals or inpatient psychiatric facilities are also large parts of agencies that do provide aftercare care

coordination, but I don't think that's the point.

I think the point is that the onus really is on the inpatient clinical team to identify the individuals who need the various levels of intensive care coordination to ensure follow up. And that I think is a key element of quality that's on the inpatient team. So I think it's very important to note that.

And I had a question for the developers as well. In terms of the readmission; I couldn't find too much data, or maybe it's there and I couldn't see it, but what percentage of the readmissions were to inpatient psychiatric facility versus general med/surg units?

**CO-CHAIR TRAVIS:** I think they're looking for it, Tom. They're busy looking.

**MEMBER SMITH:** I have a bunch of other questions that are really validity-related. We haven't voted on evidence yet. I think the evidence is very strong. There are other questions. I don't know how much people want to get into them now versus wait.

CO-CHAIR TRAVIS: We will wait, if that's okay, because we need to kind of focus on the evidence. And then we'll vote and then we'll kind of move down the line. But I think they are looking for an answer to your question. Is that something that you need right now, or can they come back with it in a few minutes when they find it?

MEMBER SMITH: No, it doesn't have any bearing on the evidence.

CO-CHAIR TRAVIS: Okay. Thank you. Well, they will continue to look.

Are there any other comments or questions around the evidence?

(No response.)

CO-CHAIR TRAVIS: Okay. If not, I think we'll go to the vote.

MS. HERRING: Voting is now open for evidence for 2860. Your choices are one, yes; two, no. And I believe we're looking for 18 votes, if I'm counting correctly.

(Voting.)

MS. HERRING: Never mind. Nineteen. And we have 18 votes for yes, one for no, so 95 percent yes, 5 percent no.

CO-CHAIR TRAVIS: Okay. You all let me know when you're ready. Are you ready?

MR. CAMPBELL: So, yes, just to make a comment that 75 percent of the readmissions are psychiatric and 25 percent of the readmissions are for medical.

CO-CHAIR TRAVIS: Okay. Thank you.

Tom, hopefully you heard that.

MEMBER SMITH: Yes, thanks.

CO-CHAIR TRAVIS: Sure. Now we're going to go to the performance gap or the opportunity for improvement. And we will start, Tom, how about with you?

MEMBER SMITH: Sure. I think from the data they present for the overall population the mean 30-day readmit was 21, around 21 percent, minimum 12 percent, max closer to 31 percent, and the 10th and 90th percentiles were 17 to 24 percent. Then in comparison to national readmit

rates they did have -- 8 percent of the hospitals were -- performed better, and 13 percent performed at the low end. So I think there's a significant gap there in opportunity.

CO-CHAIR TRAVIS: Thank you.

Frank, any additional comments?

MEMBER BRIGGS: Nothing to add.

CO-CHAIR TRAVIS: Keith?

(No response.)

CO-CHAIR TRAVIS: Okay. Any questions or comments from the group around the opportunity for improvement or gap? Yes?

MEMBER ROBERTS: So I still wonder if at the high readmission rates, especially if 75 percent of them are for psychiatric, if it's not an issue of some of our failing behavioral health issues for access versus if it's a quality. It just concerns me.

MEMBER SMITH: I think the developers can speak to that. They did look at regional measures of access and actually found; I'm trying to -- off the top of my head, higher rates of

readmission in the areas where there was greater access. It's a complicated set of circumstances.

I think one issue is access, the idea that if there are not enough providers out there that people will have higher readmission rates, but we often see the opposite. In areas where there are more providers, you see greater use of services. It's due to a couple things perhaps. One is migration. People with especially serious mental illness tend to migrate and are more likely to live in areas where there are providers.

The other is the opposite. And one of our concerns for access is in areas where there is not access to services, you see a lot of trans-institutionalization or service use. People are ending up in prisons, in jails and homeless settings. And so, readmission rates actually might go down in these rural areas.

So it's a very, very complicated conceptual and real landscape and I don't think it would be fair to say that we should not pursue

a measure like this just because we know there are significant access issues. But I'm speaking -- I'm interested in the developer's response.

CO-CHAIR TRAVIS: Yes, please.

DR. WINTERSTEIN: This is indeed a really complicated issue, and the observation is correct that in more urban areas there's actually more readmissions. And that could be a consequence of more complex issues, perhaps substance use disorder, but it could also be a consequence of more access which would essentially look like it's a quality issue, even though it's not. And we can address this as well. And we have done a lot of analysis to this.

We actually looked at access in various ways. We looked at RUCA designation. We geocoded patients. We used the HRSA health shortage area designations which produce county-level estimates of health shortage areas broken down by access to medical versus psychiatric care. And then we also created our own access

measures within the Medicare environment where we geocoded patients, basically produced a 20-mile radius circle around them, counted the number of providers that were in that circle and created ratios of that.

So basically the number of patients we had, Medicare-eligible patients with psychiatric disorders in that circle versus -- and then in ratio toward the number of providers that were there. And we did this with respect to non-prescribing mental health care providers, so essentially psychologists as well as psychiatrists, IPFs. We even looked at pharmacists and PCPs. None of them were particularly predictive of readmissions, so the access issue didn't really bubble up to the surface as being the major explanation.

Now that said, access means very different things, right? There may be providers around me and I still may not see them. And I think that is another layer that it's I think much harder to get to, so we know for example

from the literature that living alone will increase my risk for readmission because there's not that direct support network around me which might actually encourage me to see a physician and to get help. And that may not be there. So there's different layers of access that we certainly don't all capture in this particular measure.

The other thing I think that we need to be aware of is that this measure is built for Medicare patients. We are not looking at the homeless patients with no insurance. So I think the global population of IPF patients is certainly more diverse than what we have here in front of us in the Medicare population.

CO-CHAIR TRAVIS: Thank you.

Leslie?

MEMBER LESLIE HALL: Is it possible that at hospitals there's -- that is accepting more patients for readmission as the safety net hospital that these measures can potentially identify where dumping is happening,

inappropriate movement of patients to a facility? That's one question.

And then we wouldn't want to have a hospital that is accepting these patients to be penalized because they are the community provider of these services when they're acting as that main hospital readmit versus the originating hospital, so the medical facility.

DR. WINTERSTEIN: Not completely. So when you say hospitals who would accept those patients versus those who don't --

(Simultaneous speaking.)

MEMBER LESLIE HALL: I'm in a rural community and in our state there's only a few hospitals that generally accept the patient in the emergency room, I mean, whether it's a cultural bias or a medical bias or an access bias. And so they will see a disproportionate share of any mental health patients or patients who are being readmitted and don't have access to the inpatient psychiatric facility but are coming into the emergency room of the medical facility

just for access issues.

And they're the only ones accepting those patients. They will have a disproportionate number of patients simply because of who they are and how they accept it. The faith-based hospitals are very -- would that unfairly penalize them on one hand, and on the other hand would it also identify where we have a problem of hospitals that aren't fairly sharing that population?

DR. WINTERSTEIN: Yes.

MR. CAMPBELL: Maybe just one thing to clarify to start. The readmission penalty is assigned to the facility that initially accepts the patient, right? So in the cohort the first facility would be the facility that would get counted for the readmission.

MEMBER LESLIE HALL: Unless that readmission is now for a medical primary purpose.

MR. CAMPBELL: Right. Well, all-cause readmissions are in our measure, but acute care hospitals are not in the denominator of the

measure, so the initial admission --

MEMBER LESLIE HALL: Oh, that's back to your point?

MR. CAMPBELL: Yes, the initial admission has to be an IPF in the cohort.

DR. WINTERSTEIN: But perhaps to elaborate on this a little bit more, the denominator are really the IPFs, and these are certainly -- the largest proportion are inpatient psychiatric units that get paid under the IPF model. And certainly those inpatient units have the more complex patients, because if you have the choice between a freestanding IPF where the requirement for medical care is limited to a psychiatrist versus a hospital that's imbedded in an acute care facility, the patients who go in the acute care facility imbedded IPF are more likely to have more complex medical problems, which of course is incorporated in this measure. So we use the same risk-adjustment methodology with respect to all medical comorbidities that all the other readmission measures are using.

CO-CHAIR TRAVIS: Pam?

MEMBER ROBERTS: You mentioned that you included Medicare only, and that included the chronically disabled patients that are on Medicare, is that correct?

DR. WINTERSTEIN: Yes, and that's in fact the larger proportion. So this is a very different population than what we usually like to think about when we think about Medicare. So the largest proportion is below 65 years of age because they quality because of disability.

MEMBER ROBERTS: Right. Thank you.

CO-CHAIR TRAVIS: Okay. Paula?

MEMBER MINTON-FOLTZ: Have you also looked at process performance gap in say like the Joint Commission HBIPS data that would seek to -- they were processes that would seek to create maintenance so patients wouldn't have to be readmitted.

MR. CAMPBELL: Yes, so there are HBIPS measures slated for reporting in the IPFQR specifically, and also measures of care

coordination as process measures. So those will be in the data set for -- they'll be reported.

CO-CHAIR TRAVIS: Okay. Any other comments or questions around performance gap?

MEMBER MINTON-FOLTZ: I was just --

CO-CHAIR TRAVIS: Sorry.

MEMBER MINTON-FOLTZ: To follow up on your -- was there a performance gap in those measures? You said they're included, but did you see a gap?

MR. CAMPBELL: I believe there was, but I don't have the data specific to the testing of those measures and the population. I don't believe they've been reported yet.

CO-CHAIR TRAVIS: Thank you. Okay. Seeing no other comments or questions at this time, we'll go on and take a vote on performance gap.

MS. HERRING: Voting is now open for performance gap for Measure 2860. Your choices are one, high; two, moderate; three, low; four, insufficient.

(Voting.)

MS. HERRING: The results are 8 high, 10 moderate, 1 low, 0 insufficient, so 42 percent high, 53 percent moderate, 5 percent low.

CO-CHAIR TRAVIS: Okay. Thank you. Now we're going to move on to reliability.

And, Keith, anything that you want to bring up?

MEMBER LIND: Not really. I think the developers actually covered this pretty well. They did use the two methods, split the sample in half and bootstrapping. And for splitting the sample the correlation coefficient was 0.6, which they called moderate. And for bootstrapping, 0.78, which they thought was a substantial level of agreement.

CO-CHAIR TRAVIS: Thank you. Tom, any additional comments on reliability?

MEMBER SMITH: No, nothing to add.

CO-CHAIR TRAVIS: And, Frank?

MEMBER BRIGGS: Nothing to add. Thank you.

CO-CHAIR TRAVIS: Thank you. Any comments or questions from the Committee on reliability?

(No response.)

CO-CHAIR TRAVIS: Okay. I think we'll go to the vote.

MS. HERRING: Voting is now open for reliability for Measure 2860. Your choices are one, high; two, moderate; three, low; four, insufficient.

(Voting.)

MS. HERRING: We have 3 high, 16 moderate, 0 low, 0 insufficient, so we have 16 percent high, 84 percent moderate.

CO-CHAIR TRAVIS: Great. Okay. We'll go to validity.

Frank, any comments you want to make on validity?

MEMBER BRIGGS: Yes. So for validity they did face validation with their TEP, across their TEP of 17 members. The median rating was seven, which they described as agreement.

Looking at agreement versus non-agreement, it was a 60/40 split, so 60 percent had the rating of 7 or 9, which they said was agreement, neutral at 6 votes, and disagreement was 1 vote.

In terms of threats to validity looking at the patients that were excluded, the majority of the patients that were excluded came from transfers or interrupted stops. I suspect that's the patients who may have been transferred from one unit to a medical or vice-versa. That accounted for 7.2 percent.

And then they addressed the discharge against medical advice, which was 1.2 percent of their sample at 9,000 patients leaving against medical advice.

The did look at SDS adjustment. Ultimately did not include it in their recommendation. And I think that's it.

CO-CHAIR TRAVIS: Okay. Thank you so much.

Tom, any comments on validity?

MEMBER SMITH: Sure. I have a couple.

Yes, they did a fairly comprehensive SDS analysis. As in with previous measures the variables just didn't hold up in the end, so they ended up proposing not including them. I don't know whether to be surprised or not. I guess I'm not surprised given the complex interrelations between behavioral healths and SDS, but the effect size -- or the odds ratios were washed out when controlled for clinical variables.

I have questions about a couple of things. The interrupted stays and the age issue. As regards the interrupted stays, because of the CMS billing procedures they're not able to count any readmissions that took place on the day of or day one, day two following discharge. And I don't know any other way around that. I don't know Medicare claims as well as I know Medicaid, but from my understanding there's no away around that. But I do worry as a threat to validity that this excludes a fair number of people who are discharged from an inpatient psychiatric facility and then readmitted a day or two later.

A lot of people are in fact readmitted.

One common example is to include people who are still acutely suicidal and get home and it's another crisis and they have to go back, or people who have unstable housing environments and they just don't connect with the aftercare housing arrangement. And they're back the next day or on day two.

There's a lot of such people, the Medicare claims however just simply don't allow for a meaningful way of capturing them. And I'd like to hear the developers -- if there's any further thoughts or ways they think this might be addressed in the future.

I also have a question I just -- my second issue is about the age, in the 65 and older population versus the younger. The measure does group all individuals aged 18 and up. Out there in the clinical world there tends to be a distinction between geriatric psychiatry programs versus general adult psychiatry programs. And it varies around the country depending on access and

geography, but a lot of clinician experts would say that geriatric psychiatry is somewhat different from adult psychiatry. You see much higher rates of course of dementia disorders in that elderly population. And I think that was born out in the data.

I think the developers discussed this with your expert panel and you did a series of cohort analyses breaking out by dementia and by age and you found that the models were not as robust. The c-statistics were actually lower when we stratified your cohorts and broke out your models. And so I understand that. I guess the best model from a statistical perspective is to include all age groups, but from a clinical perspective I think a lot of providers would say that the geriatric psych population is some ways fundamentally different from the younger population. So I wonder if you have any other thoughts about that.

MR. CAMPBELL: So just the first issue with regard to the interrupted stays. So that is

a situation that's unique to the inpatient psychiatric facility billing. And we worked very closely with the CMS payment folks to even understand that and how the claims were collapsed. So we were unable to count those in the data, but it's definitely something that as we move forward could definitely be something that we would work on in the future.

And then I'll defer to Almut on the second question you had.

DR. WINTERSTEIN: There's one thing to add about the transfers. One thing is that all the readmission measures remove the first day post-discharge from the readmission time frame because of the transfers. So the people we lose in addition because of this weird readmission model is really just one day. That would be that day two because this still is the issue that those readmissions would still be reimbursed under the same original claim. So there's not two claims that get generated.

The other part in terms of losing a

good proportion because of that, this is actually -- I think we have in our report a histogram that shows the number of readmissions over the 30-day time frame, and that looks quite different from the acute care readmission frame work where there is a much steeper decline in the readmission rates over the 30-day period.

In our case I wouldn't say it's flat, but the decline is much, much slower. And we actually followed this through 90 days. And you see we have here an average readmission rate of about 22 percent. If you go to 90 days, we have 50 percent. So there is a consistent readmission that goes across the entire time span. If you wait six months, it still increases quite a bit. So it's not that we are losing a whole lot proportionally of all the readmissions if we are dropping that additional day.

With respect to age, we did two things actually. One was that we did do, as you mentioned, stratifications by age and created separate risk-adjustment models for different age

groups. We also stratified by different types of diagnosis, specifically Alzheimer/dementia patients versus the more streamlined mental health disorders. So cognitive disorders was versus mental health disorders basically. And as you said, the c-statistics were actually lower than combining everything, which of course relates to just getting more explanatory power from the larger sample.

The other thing that we did; and that's probably a little bit more of an unusual way that you may not see in previous readmissions measures, we actually we did a multinomial model where we were not predicting readmission versus not, but we actually predicted readmission for a major -- with a principal diagnosis for a mental health disorder versus a physical disorder just to also touch more on patients who might have more multimorbid issues, which would be obviously more in the older population.

And this modeling approach, which is way more complex and took about two weeks or so

to run on a very high-speed server, still didn't really add a whole lot. And my explanation to this is that by capturing the multimorbidities -- which I think we did fairly well, in particular because we didn't use the typical classifications that are used for the mental disorder. We broke them further down to really get enough granularity. I think we captured the age issues primarily really with the clinical conditions so well that it didn't really require anything in addition.

MR. CAMPBELL: And one additional comment, if I could make, related to the readmission patterns, which I think is really important for this population, is that one-third of the population is readmitted more than one time, and five percent of the population is readmitted five or more times. So within this population readmissions are a significant issue.

CO-CHAIR TRAVIS: Thank you for that.

Keith, have I give you a chance to say anything?

MEMBER LIND: You did.

CO-CHAIR TRAVIS: Okay. You're fine?

(No response.)

CO-CHAIR TRAVIS: Okay. That sounds good.

So I know Bruce had his card up early, so if you'd like to make a comment?

MEMBER BRUCE HALL: Yes, two comments on validity or model choices, or maybe bordering into use, but the first is with respect to the expert panel face validity query, 7 out of 17 of your preferentially selected experts weren't impressed with face validity. They were either neutral or negative. That's a little more than you would expect I think in a preferentially selected group of experts, so that raised a flag for me.

The other issue I wanted to ask about is why you're setting the performance here at 24 months. I know the standard in your material said that you did it for stability or power, but I'm not sure that we actually know enough about

what the institutional grades will look like to know that you need two years for performance. Certainly you could always use two years of data for derivation, but do you really need two years of data for performance?

Because I think probably most of us in the room in our normal lives we don't like measures that stretch on for extended periods. They became quite hard to interpret and make use of and drive improvement off of.

So I don't think we considered any other measure during our two days that was longer than a year and I was wondering why that decision was made. So those were the two queries I had.

MR. CAMPBELL: Those are two are really good questions, and maybe start with the latter question and then go to the former.

So one of our goals here was -- again this program is a pay-for-reporting program and not a pay-for-performance program like the other readmission reduction program. And we wanted to have as many IPF facilities that would meet that

minimum threshold of 25 cases, which has been the standard for the CMS measures. So by using the two-year time frame and accumulating those cases, we get about 96 percent of the IP units in the country that could be reported using the measure. So that's the rationale for that.

As far as the face validity vote from the panel, some of our neutral panel members were just that they felt that they didn't have enough expertise to assess this particular measure. And then we also had some concern about the measure being all-cause versus psychiatric readmissions. And we wrestled with that quite a bit, but we think that we're taking a very patient-centered approach and that all readmissions are very intrusive to patients and burdensome and burdensome to facilities.

And we also think it's very difficult to tease out the rationale for why a patient gets readmitted, so for example one of the most common -- or one of the more common readmission reasons was adverse drug events. So those adverse drug

events for psychotropic drug poisoning show up in the data. So we just didn't feel that it made sense to separate those things out.

So I think that's reflective of what you're seeing in the vote, but the folks that did support it supported it very strongly. And in our public comment period when we took the measure out for public comment, 83 percent of stakeholders supported the measure as you see it.

CO-CHAIR TRAVIS: Thank you.

Pam?

MEMBER ROBERTS: Just a quick question. Did you look at the readmissions by freestanding versus unit? And if so, was there a difference?

MR. CAMPBELL: We did, and there were no meaningful differences between the two.

CO-CHAIR TRAVIS: Kathy, is that your card?

MEMBER AUGER: So this question may be reflective of my ignorance as a pediatrician, so I'm sorry. You mentioned earlier that the

majority of the patients in this measure are in that 18 to 64 range, but my question really is what percentage of people with significant mental illness end up on Medicare in the first place? So how much does this type of measure reflect the vast -- like how hospitals will perform with the majority of their patients as opposed to just the Medicare, if that makes sense.

MR. CAMPBELL: I think that's a really good question and one I think we may have to get back to you on, because we only have data on the Medicare population, from the fee-for-service population for these.

MEMBER AUGER: And but do you have any sense of what percentage of people with significant mental illness would have Medicare? Is it --

(Simultaneous speaking.)

MEMBER SMITH: It's Tom. I'm echoing. I could jump in.

CO-CHAIR TRAVIS: Okay.

MEMBER SMITH: It varies from state,

but if you look at state-level populations of individuals with serious mental illness, and you see the majority of them will be on Medicaid, the number that are duals, like it varies from state to state, but I think on the low end it's 15 to 20 percent. And then I think it goes up from there. Might be 30 to 40 percent. I can't exactly remember, but it's on that order. It's a significant sub-population.

CO-CHAIR TRAVIS: Thank you. Okay. Wes?

MEMBER FIELDS: Yes, it was actually a very good question, because in order for somebody with a chronic behavioral illness to become Medicare-eligible they probably have to have become unstable. Schizophrenia is the classic sort of benchmark. Bipolar disorder is very similar.

But the reason I wanted to speak to this is that it's great that you have a pay-for-performance measure, but in terms of community needs and crises in terms of access to behavioral

health services, this is the wrong piece of the pie. Because the real issue are young adults and late adolescents who are becoming unstable before they're diagnosis or before they're actually stabilized in an inpatient setting. Classically the first time they get admitted either they're insured by their parents or they're young adults without coverage at all, depending on where they live, state by state.

As Tom suggested, there's wide variation between states, partly with the Obamacare effect, about whether or not they get Medicaid benefits, but really the meter starts for when they become Medicare-eligible, when they have a significant psychiatric disorder that results in their permanent disability. So classically these patients become impaired, disabled. It persists long enough for them to wind up with Social Security benefits that are more or less tied to Medicaid.

But the Medicare benefit is sort of like the last thing that happens to them in terms

of coverage and benefits, so it's great that you're looking at persons greater than 18 years old with Medicare, but the real systemic problem of access are people who are not yet Medicare-eligible and often not yet Medicaid-eligible who are at great risk in the community in emergency departments where they can hang around for days at a time.

So I guess I'm speaking in favor of future measures, but I'm sure you both know, as Tom knows, that the real issue here are people that have very unstable behavioral health problems who have not been sick long enough to become Medicare-eligible.

MR. CAMPBELL: Thank you. Thank you for those comments. And the only one thing I would clarify is it's pay-for-reporting, not pay-for-performance. So in this case we're not looking at performance. But thank you.

DR. WINTERSTEIN: But he's happy to say it again.

(Laughter.)

DR. WINTERSTEIN: I mean, there is research that even shows that mental health disorders also correlate with the ability to seek insurance, so Medicaid for example -- well Medicare as well, but both require us to sign up for it. And I am paranoid and I don't want to deal with the system, I may very well have no insurance. So the problem is way more complex, for sure.

CO-CHAIR TRAVIS: Thank you Okay. Last call for any questions or comments on validity.

(No response.)

CO-CHAIR TRAVIS: Okay. Let's go to a vote.

MS. HERRING: Voting is now open for validity for Measure 2860. Your choices are one, moderate; two, low; three, insufficient.

(Voting.)

MS. HERRING: The results are 16 moderate, 4 low, 0 insufficient, so 80 percent, 20 percent low.

CO-CHAIR TRAVIS: Okay. Now we're at feasibility.

Tom, any issues relative to feasibility?

MEMBER SMITH: No, I think the developers described it well. These are claims-based and all the elements have been -- may of the elements for the risk modeling have been validated, so I think they're quite feasible.

CO-CHAIR TRAVIS: Thank you.

Keith, any other comments?

MEMBER LIND: No, I agree.

CO-CHAIR TRAVIS: Agreed? Okay. Frank?

MEMBER BRIGGS: No, I agree.

CO-CHAIR TRAVIS: Okay. Thank you. Any other questions or comments from the Committee on feasibility?

(No response.)

CO-CHAIR TRAVIS: Okay. Let's go to a vote.

MS. HERRING: Voting is open for

feasibility on Measure 2860. Your choices are one, high; two, moderate; three, low; four, insufficient.

(Voting.)

MS. HERRING: Just waiting on -- oh, never mind. The results are 12 high, 8 moderate, 0 low, 0 insufficient, so 60 percent high, 40 percent moderate.

CO-CHAIR TRAVIS: Okay. And our last category other than overall, we'll look at use and usability.

Frank, any comments?

MEMBER BRIGGS: Although the program is not currently publicly reported, it has been submitted for inclusion in the CMS Inpatient Facility Quality Reporting Program.

CO-CHAIR TRAVIS: Keith, anything to add?

No audible response.)

CO-CHAIR TRAVIS: Tom?

MEMBER SMITH: No, I think it's a very important measure. We don't have anything like

it. We've got very little in behavioral health. As you can see, there's a lot of validity issues, a lot of conceptual issues that have to be sorted out, but we need these measures. Especially as pay-for-reporting I think there's going to be a lot of interest and a lot of potential use.

CO-CHAIR TRAVIS: Thank you, Tom.

Paul?

MEMBER HEIDENREICH: In terms of potential unintended consequences is there any other outcome measure for this group such as mortality or suicide? Because I think especially if we're talking about for young people, I know within the VA we have the issue of not being able to track them and losing them to follow up, and then they end up incarcerated or dead. And I think that's probably less of an issue for the Medicare population, but I'm just curious, is there any other part? Like MI, heart failure, pneumonia we have a mortality measure.

MR. CAMPBELL: No, currently there aren't any mortality measures for this

population.

CO-CHAIR TRAVIS: Any other questions, comments or others on use and usability?

(No response.)

CO-CHAIR TRAVIS: Okay. We'll go to the vote.

MS. HERRING: Voting is now open for usability and use on Measure 2860. Your choices are one, high; two, moderate; three, low; four, insufficient.

(Voting.)

MS. HERRING: And the votes are 6 high, 13 moderate, 1 low, 0 insufficient, so 30 percent high, 60 percent moderate, 5 percent low.

CO-CHAIR TRAVIS: Okay. Any comments before we vote on overall suitability?

Yes, Carol?

MEMBER RAPHAEL: I just wanted to echo what Paul said, because to me I'm very much in favor of moving ahead. I think this is really important directionally, but I don't think this is classic kind of transition. I can remember

that for the 30 percent of referrals we got with a psych diagnosis, 30 percent of those referrals we could not connect with the person. And it was a combination of not being able to find the person, being able to find the person, but the person not wanting to engage and distrusting, a variety of factors. But I just think that as we think this through we should be cognizant that just there's an engagement that's really, really important with this population. It isn't just a treatment, an intervention.

CO-CHAIR TRAVIS: Thank you, Carol.

Any other -- Paula, is your card still up for a comment?

(No response.)

CO-CHAIR TRAVIS: Okay. Any others?

(No response.)

CO-CHAIR TRAVIS: Okay. Well, we'll go on and vote on overall suitability for endorsement.

MS. HERRING: Voting is now open for overall suitability for endorsement for Measure

2860. Your choices are one, yes; two, no.

(Voting.)

MS. HERRING: The results are 19 yes, one no, so 95 percent yes, 5 percent no. Thank you.

CO-CHAIR TRAVIS: Okay. Thank you very much for that. We are going to go to a time of public comment. And I will first open it up to those in the room. Any public comment in the room?

(No response.)

CO-CHAIR TRAVIS: Okay. Seeing none. Operator, could you please open up the lines for public comment?

OPERATOR: Yes, ma'am. At this time if you'd like to make a comment, please press star then the number one.

(No response.)

OPERATOR: There are no public comments at this time.

CO-CHAIR TRAVIS: Okay. Thank you, operator.

I'm looking at -- are we going to continue or do we go?

MS. O'ROURKE: So lunch is ready and we haven't had a break yet, so why don't we take a 15-minute break and then come back for a working lunch to discuss our next measure? So we'll see you back here shortly before 1:00.

CO-CHAIR TRAVIS: Okay. Thank you, all.

(Whereupon, the above-entitled matter went off the record at 12:40 p.m. and resumed at 1:00 p.m.)

MR. AMIN: All right, if everyone can find their way back to the table, we're going to get started here. Thank you.

CO-CHAIR BULGER: So the first measure here after lunch is 2884, 30-Day Unplanned Readmissions for Cancer Patients from the Seattle Cancer Care Alliance and the Alliance of Dedicated Canter Centers. Thanks for being here.

So as we have in the past, we'll start with the developers. Thank you.

MS. JAGELS: Good afternoon, and thank you for having us. I'm Barb Jagels. I'm Vice President of Quality, Safety, and Value at the Seattle Cancer Care Alliance. Joining me this afternoon in person is Terry Fisher, our project manager based at MD Anderson. On the phone with us today we have Susan White at the James-Ohio State who served as our statistician; Denise Morse, a program manager who greatly assisted with the development of this measure based at City of Hope.

So it is with a great deal of professed anxiety and humility that I sit with you today. It's been a very sobering experience to hear you all this morning and have the opportunity to sit with you and see how this work is actually done.

We came last year to the MAP where we presented our measure, received the feedback and went back and obviously sustained a great deal of further testing and learning. So I sit with you today having seen our Yale colleagues in action

this morning and recognizing that our measure looks very distinct and different, so I look forward to learning from your input and wisdom this afternoon.

So I'll start briefly by describing, of course, that we are a unique group in that we're 11 PPS-exempt cancer centers. What we were finding three years ago is that all-cause readmission measure wasn't working well for our cancer purposes. Naturally, you'd observe that cancer patients enter and exit the hospital regularly, sometimes for intentional, appropriate reasons, think chemotherapy and surgery; sometimes for foreseeable and avoidable reasons, think chemotherapy and nausea.

So as we started to look at our own data and collaborate across our centers, we asked ourselves the question could a group of team spirited, quality improvement minded hospitals work together to share our data, compare our claims, clean up our coding, and most importantly understand where we have opportunities to better

serve our patients, specifically related to foreseeable and avoidable cancer treatment related symptoms?

So I'm delighted to tell you that we think we've made great progress and if we have time or interest this afternoon, Denise and I can share with you specific narrative examples where we've used this data to actually improve our patients' experience of cancer, but I know that's not our intent this afternoon. We're here to explain to you how we did our testing and have you give us input on how we might proceeding.

So with that, I'll open it up for questions.

**CO-CHAIR BULGER:** Okay, so we'll start with -- as we have, with the evidence. And we'll start with Helen.

**MEMBER CHEN:** Thank you, this is actually an interesting measure and as they mentioned, it's a very narrow focus. And I would just ask a question as to whether it's appropriately named. It's a measure that looks

at cancer patients cared for specifically at PCHs. So I don't know if we want to be more specific because I don't -- that was just my first question. But in terms of the evidence that was provided by the developers, they did cite some relevant literature regarding the importance of condition-specific or disease-specific measures for quality that were above and beyond the typical evidence that we've seen to date regarding the ability of facilities to intervene on readmissions through general quality improvement measures. So there is some evidence implemented.

CO-CHAIR BULGER: Keith.

MEMBER LIND: It seems like an important measure. I agree that it's important, but I defer to your comments. I agree.

CO-CHAIR BULGER: Cristie.

CO-CHAIR TRAVIS: No, nothing to add. Thank you.

CO-CHAIR BULGER: Okay, so I think we'll come back to that. But I think we'll vote

on the evidence standpoint so it's yes or no.

MS. HERRING: Voting is now open for evidence on measure 2884. Your choices are 1, yes; 2, no.

(Voting.)

MS. HERRING: And it's 17, yes; 0, no. So 100 percent yes.

CO-CHAIR BULGER: Okay, great. So go to the gap.

MEMBER CHEN: So in terms of the performance gap, the developers report interestingly that the performance on readmission rate for various groups of hospitals, the first group being the ADCC and then a second group, the Coalition for Quality Improvement which is actually, there's an overlap between those two groups, as well as some additional data from the bigger UHC group, larger group of 100 hospitals.

And the rates of readmission that were reported in the alpha group, there's two different time periods, was 13 to 13.4 percent. And in the beta group which was a year later, was

14.5 to 15.8 percent. Interestingly enough, in the validations that they report on readmission rates, unadjusted readmission rates for six hospitals in the ADCC group and there does appear to be some degree of variability in performance. And the rates for those hospitals were, let's see -- I don't have it. There's a range, 10.9 to, I believe, 15.7. I may be misquoting, but there is a performance gap.

MR. AMIN: Helen, would you mind just moving your microphone closer to you. There's construction over here to the right. It's making it difficult to hear.

MEMBER CHEN: Sure, sorry.

MR. AMIN: Thank you.

CO-CHAIR BULGER: Keith.

MEMBER LIND: Nothing to add.

CO-CHAIR BULGER: Okay, anything from the committee on the performance gap?

MEMBER CHEN: Did you want me to talk about disparities here too, as well?

CO-CHAIR BULGER: Say that one more

time?

MEMBER CHEN: I'm sorry, should I talk about disparities here as well? Okay, so they did take a look at adding some SDS factors to their risk model and they were, of course, limited by data points that were available on electronic claims data. Race and gender were considered as risk adjusters, but were not felt to be significant.

Interestingly enough, they decided on payer class as a marker for SES and the payer class lumping -- I'm not a lumper, but they appear to be lumpers. And it was basically Medicaid, charity care, and no insurance.

And I'm wondering from a committee perspective we really feel that that's a valid delineation of SES as a marker, although in the patients who had the SES composite factor, versus who didn't, there actually was a higher readmission rate. I don't know if the developers want to speak to that.

My sort of anecdotal experience of

regional cancer center care is that some proportion of the uninsured are actually people who are private pay who have come to a regional center, but looking at your overall statistics, it looks like that was only about 3, 3.5 percent of the total.

MS. JAGELS: Susan, can you take that one?

DR. WHITE: Sure, So I guess we are lumpers. I never thought of it that way. So yes, you're correct. Depending on which of the -- there's a small number of centers and depending on which ones they are, some of them have more or less international pull, which would end up -- be out-of-pocket payers mostly.

And so what we're trying to do is come up with obviously a proxy that would not have such small -- that had a reasonable number of observations so that we could really assess the it versus sort of as a proxy. It would be better if we had some geographic indicators or something else perhaps, but we tried to do the best we

could with what we had available and we did see a difference. So I felt as though sort of lumping them together might dampen out the cash-paying patients that we might see in the traditional self-pay for this particular type facility.

CO-CHAIR BULGER: Wes.

MEMBER FIELDS: Real quick, I think you'd be doing Medicaid beneficiaries a real favor if they suffer from cancer by looking at that as a separate category from the no-pay, self-pay patient, especially if you've got international patients at major New York type centers. My experience in multiple hospitals over a long period of time is that Medicaid beneficiaries have delays in getting staging done. They may or may not get primary surgery done. They may or may not get chemo, but I think you're very likely to see substantially higher readmit rates that are unscheduled because of the gaps in their community care and so you'd be doing them a real favor if you could separately track that from the self-pay patient.

MS. JAGELS: I think that's good advice.

Susan, anything you'd add in response?

DR. WHITE: No. I think that we can certainly adjust and look at it that way.

CO-CHAIR BULGER: Cristie?

CO-CHAIR TRAVIS: I want to take us back just for a moment to be sure that I kind of understand the construct of the measures so that I can think about the performance gap.

When I read this, it looks like it's readmissions back to the same hospital. And I guess the only thing that -- I don't know it's the only thing, I shouldn't have said it that way. I guess one of the things that would concern me about that is that there could be potentially admissions to another hospital for cancer care, thinking that some people may be traveling quite a distance to go to one of your 11 hospitals. And so this measure doesn't really capture that. And so I think -- and it's not the traditional way we've been looking at readmission

measures over in the general acute care hospital setting, so just trying to figure out perhaps who is not being captured, but also thinking about the construct from the beginning in terms of why it was set up that way.

MS. JAGELS: That's a really great question and indeed, we've had copious amounts of internal debate about the data we don't have. So we don't have CMS data, nor do we have anything that looks proximate to an all-payer database. What we have is our own each individual's hospital claims data.

So indeed, we did not set out to test the hypothesis that we were failing to coordinate care beyond our hospitals. You're absolutely right. And many of our patients do come and go within our regions, or even nationally, so we recognize that there are some circumstances in which this measure doesn't tell us all that we need to know about whether we're successfully managing those patients outside the hospital settings outside of our index hospital area.

Instead, what we've really focused on is given the patients we can find and given the patients who are going in and out of the hospital regularly, what sorts of clinical circumstances are we seeing? So as an example, I can't tell you in Seattle who goes back to Wenatchee and gets -- in Central Washington -- who gets admitted there. I can tell you that we draw a lot of patients from Central Washington to Seattle for leukemia and lymphoma care. And I can see why they're going in and out of the hospital particularly for foreseeable and preventable reasons.

So this measure has limitations. I think it's useful for our academic cancer centers. It allows us to compare. And I think fairly rigorously so to understand what the circumstances are and is helping us actually develop triggers. It doesn't help us broadly in the care coordination issue.

CO-CHAIR TRAVIS: Well, thank you and I think especially since this measure has been

looked at by the MAP. And if I'm remembering correctly and that it's coming here for endorsement, I think that this distinction and this difference than most readmission measures is an important piece for us to be sure that it gets translated as it goes through any other related processes here at NQF. So thank you for that.

CO-CHAIR BULGER: Helen?

MEMBER CHEN: I just wanted to add that maybe changing -- sorry to harp on this -- the name of the measure might help that because it's not really a general measure of care of cancer patients. It's really patients who are cared for at those PCH hospitals during their index hospitalization. Just a thought.

CO-CHAIR BULGER: Any other -- Leslie.

MEMBER LESLIE HALL: So is it a very high percentage of patients who have diarrhea, nausea, dehydration during care that there's often a readmission or an observation or something? Is it a high percentage that makes this difficult to track as a result of these

distances that we talked about earlier of patients traveling, not going to travel four hours if they're in the middle of dehydration and diarrhea. They're going. And yet, is it prevalent? Is it really common?

MS. JAGELS: About 13 to 15 percent of the time, according to our findings. What we were really trying to set out to establish is that there was so much white noise in the all-cause readmission that we couldn't see the underlying patterns. There were so many patients intentionally going back into the hospital for scheduled surgery or for scheduled chemotherapy that when we went looking for the things that they ostensibly shouldn't or less ideally be admitted to the hospital for, we couldn't find them. So once we excluded those intentions, then indeed -- there were three: nausea, vomiting, diarrhea, gastrointestinal side effects, pain, and septicemia.

So now septicemia is a little trickier to consider foreseeable and avoidable except in

the circumstance of high-intensity chemotherapy or other treatment. Nausea, vomiting, diarrhea, and pain eminently foreseeable and avoidable. We want to do a better job of getting ahead of those circumstances and preventing those admissions. So it was less around the travel dynamic and more around we're doing things that we know will cause these side effects. Let's develop mitigation strategies to once again more successfully deliver that therapy outside the hospital setting.

CO-CHAIR BULGER: Wes.

MEMBER FIELDS: Yes, real quick. I'd be a little bit concerned if the index is an admission to the hospital that this may not be a sensitive at 30 days as it would be with a longer period of time trailing the index alternately if the index could be a schedule infusion or a chemo session. I think you'd be in better shape. Because my sense of this, and I've been on both sides of this as a caregiver, that a lot of these things happen further downstream than 30 days

from the index.

MS. JAGELS: I think you're right. Denise, do you want to speak to that?

MS. MORSE: I believe what we've seen within the data is that our average time to readmission tends to hover between 10 and 12 days for the medical oncology population and when we get to the hematology it can out a little bit further, closer to 20 days. That's where they all tend to cluster. That's including up to 90 days of readmissions.

MS. JAGELS: I think we did most of them. Thank you, Denise.

CO-CHAIR BULGER: Karen.

MEMBER JOYNT: I'm not sure if this is veering into the validity, but it does seem like it would be helpful to know of the different hospital readmissions what the pattern of those is. There's been prior work in pediatrics and in heart failure demonstrating that there may be specific types of hospitals for which those non-same site readmissions are more prevalent and if

you're going to use it as a -- I totally understand for quality improvement, feedback to the same hospital what's your readmission rate that that may not be a big deal if you're comparing hospitals to each other and you're missing 20 percent for one hospital, 5 percent for another, and 40 for another, it becomes a pretty invalid comparator. So if you had data that could test that at least in some subset and then reassure us, that would be helpful.

MS. JAGELS: I agree. Susan, anything you'd add?

DR. WHITE: No, other than at the time of development we certainly didn't have the data to be able to do that. And it did obviously exist. I mean our friends at Yale have it. We didn't have a data set to be able to measure that. It's a great idea and I totally understand that it's an uneven playing field if you compare -- different centers have different referral patterns, you know, and so -- agree. If we're able to obtain the data, that's something we can

address.

CO-CHAIR BULGER: Other comments? Okay. So we will vote on the performance gap.

MS. HERRING: Voting is now open for performance gap on Measure 2884. Your choices are 1, high, 2, moderate, 3, low, 4, insufficient.

(Voting.)

MS. HERRING: The results are one high, 16 moderate, zero low, one insufficient. So 6 percent high, 89 percent moderate, 6 percent insufficient.

CO-CHAIR BULGER: Okay, reliability. And we'll start with Keith.

MEMBER LIND: So the reliability testing, they did a sample chart review and they checked for -- at the facility level, they found a range of agreement of .08 to 1. The average was .77, which sounds pretty good, but I just wondered what happened. I wondered, too -- I mean you do go through potential sources of differences in definition of the planned versus

unplanned, but it just sounds like one of the hospitals was way off the chart in that.

MS. JAGELS: We would agree. Susan?

DR. WHITE: I can help with that. So one of the pieces of the logic in the numerator is looking at whether it's an emergent, elective, or urgent admission. And we did have one facility that had a different pattern in how they coded that particular data element when submitting their claims.

And we were able to sort of break it apart, so both of those facilities, the one that was at 8 percent and the one that was -- just shy of 43, both had similar but different ways of determining that urgent/emergent versus elective admission.

So we were able to trace it back to that one particular data element. So have a little bit of a reliability issue in that that data element isn't typically used as a payment indicator and so it's difficult sometimes to leverage some of those for secondary use. So

that was really the source of that variation.

MEMBER LIND: So the developer gives it a reliability low, but I guess the question is whether -- if you've gotten these other -- how are these hospitals, the one or two -- on the same page?

MS. O'ROURKE: So just to jump in to clarify, the preliminary reading is staff.

MEMBER LIND: Oh, staff.

MS. O'ROURKE: So that's from NQF staff and --

MEMBER LIND: Oh, okay. All right. So the staff gives it to them.

MS. O'ROURKE: -- it's not non-binding. It's just where -- as we saw the information using our reliability algorithm, where we would come out, but it's to the committee to make that determination. But just to clarify, that's staff opinion, not the developers.

MEMBER LIND: So I'm --

DR. WHITE: Yes, I didn't think we tagged it low. Thank you.

MEMBER LIND: I'm not quite sure how to deal with that myself. I mean, one outlier, the question is are you able to get them in line and get everybody on the same page, and do we have to consider the way they present it or the way they might be able to fix it? I'm not sure what the answer is.

DR. WHITE: Yes, I think that's a really good question. So I think given that, you know, we have a fairly captive audience of 11 hospitals, I think we can get alignment and get everybody reporting in a reliable way.

We didn't -- we obviously left those in there because if this measure were to go more broad, I think there might be more noise because of that one particular data element. And so we may want to think about a different proxy or some other way of detecting -- you know like for emergent admission, maybe we'd look for an ED ref code or something, some other way, if this measure were to go more broad.

I think for the PCHs, though, we can

get everybody on the same page. I happen to be at the center that has the 43 percent and I know we've already tightened that up. So I think it's manageable mainly because we have such a small number of providers. Is that helpful?

MS. JAGELS: Yes. Thank you, Susan.

CO-CHAIR BULGER: Paul?

MEMBER HEIDENREICH: So just to be clear, so for what we're considering is a measure for 11 hospitals, or are we considering a measure beyond 11 hospitals? Are those 11 going to -- is that clear in the measure specification if that's the case?

MS. JAGELS: So -- correct. This is a measure we put forward to be incorporated into our PPS-Exempt Cancer Hospital Reporting Quality Program administered by CMS by virtue of the ACA. We eleven centers by virtue of our exemption are required to undergo our own quality reporting, so pending your decision today, we'd be submitting this to CMS for consideration for the MUC List for next year.

So it'd be just us for now with obviously the potential down the road with additional testing and expansion, I suppose.

CO-CHAIR BULGER: So this -- it's a carved-out program that CMS has for these cancer centers, correct?

MS. JAGELS: Yes.

CO-CHAIR BULGER: Yet, it's a CMS program, which means for CMS to use the measure it needs to be NQF approved. Is that --

MS. JAGELS: That's correct.

MEMBER HEIDENREICH: They don't have to have NQF endorsement. I think they'd like to have it, right? But it's not required.

MS. JAGELS: They really like to have it, they told us.

CO-CHAIR BULGER: But that's why you brought it was because ideally --

MS. JAGELS: It was at their encouragement.

CO-CHAIR BULGER: -- you'd like to have an endorsement of the measure and it's this

carved-out set of -- okay. Helen?

MEMBER CHEN: I had a question to clarify the numerator definition, especially regarding the exclusions.

The first question, it was interesting that you use the UB-04 as the designation for unplanned hospitalizations. And I don't know that much about UBs and whether or not there's some sort of national error rate published about completion of that. That's my first question.

Second question is one of your major exclusions is progression of disease. What you're defining as a diagnosis and you gave a code set of metastatic disease, and I guess the question is, is that in a new diagnosis of metastatic disease on the subsequent readmission, and what happens if there was already an admitting diagnosis of metastatic disease on the index? How would you then judge whether the person had progression?

MS. JAGELS: That's a great question. I'd just like to point out that our first

excursion into proof of concept for this measure was ensuring amongst ourselves that we could find our cancer patients. As you can appreciate, most of our hospitals have patients in them other than cancer patients, so that was one of the proxies we put forth.

Susan, can you explain in more detailed fashion? I hope you're still there.

DR. WHITE: Which part of the question? I'm sorry.

MS. JAGELS: First the UB question, and then second, the metastatic and codes associated.

DR. WHITE: Oh, yes. So the clinical folks involved in the measure were really looking at the metastatic and if it were a new diagnosis of metastatic, it would exclude them from the numerator as we've written the measure.

Using the UB, I'm not familiar with any literature on those particular data elements. I can tell you, and probably most people who are experiencing using the UB billing data for secondary uses, that we generally consider

variables that are sort of payment determinant to be more reliable than those that are not involved in determining a payment.

So we also -- we want to make sure we're careful on how we use variables that are on the UB and are there for more tracking than for payment.

So I think this is a learning experience. I think we thought that would be a pretty solid variable, and we thought going a priori that we would have an issue with the difference between urgent and emergent and that's why we sort of put those two categories together. Although I think certainly the majority of our facilities were all coding in the same way and looking at or at least reporting that variable in the same way.

But I'm not aware of any literature nor was able to find any after we found this issue. It might be a good study to do.

CO-CHAIR BULGER: Other questions? Paul?

**MEMBER HEIDENREICH: Just quickly, is the reliability a must pass? And so, judging from -- obviously staff is not the final word, but they gave it a low and if this committee voted low, that's the end of it, right?**

**CO-CHAIR BULGER: Helen, do you still have a -- yes.**

**Any other thoughts on reliability? You know, I do want -- because I think it's validity, but I think it's part of this whole issue of same hospital. I think probably it has to impact reliability a lot as well and given the fact that there's going to be differences and there could be shifts based on -- you know because a lot of it is market share, too.**

**So I mean if you're in a situation where there's a lot of other options, you know, say like an urban area, you may have a lot of non-same hospital readmissions versus a rural area and that could change over time. So if you have a hospital closure or things like that, it's going to change that ratio which would inherently**

change the reliability for that place, but because hopefully everybody else is compared for every place.

MS. JAGELS: That's a really good point. What we tend to find, and it's difficult to measure, is that cancer patients while undergoing treatment tend to go to the hospital that their physician recommends. So assuming that they're within the same city, we didn't look, because we obviously didn't have access to that data, but our impressions are clinically that it's a closed system.

Most of our patients stay at the hospital where either they'll get a portion of their treatment in the in-patient setting or where their oncologist has admitting privileges. So I agree, we don't know what we don't know. But I do believe that most of us -- most of our hospitals tend to see the patients going in and out of the same index hospital for the purpose of their cancer treatment.

CO-CHAIR TRAVIS: I, too, don't know

whether I'm in reliability or validity. But I guess just as a matter of clarification, have you all -- have your set of hospitals been getting at all the readmission measure feedback from Medicare, from CMS?

You know, I know you were saying this is to kind of address the unique aspects of cancer care, but have you been getting that information so that you know whether or not patients are being admitted to other hospitals more broadly?

MS. MORSE: Yes.

MS. JAGELS: Yes and no. Denise, you want to take that one? Go ahead.

MS. MORSE: Yes, absolutely. We received the CMS dry run report from -- that is posted on QualityNet, that is the readmissions all-cause dry run that they do produce for us and they just do not submit that back.

We actually did, as part of the initial alpha testing for this measure, compare some of our results we identified from our dry

run to see how much we are seeing of that readmission to the same hospital and readmissions to different hospitals.

And we actually do not see as much readmissions to other hospitals. To echo what she was saying, we tend to see our patients come back to our own hospital. And so I only see a difference of about 3 percent when I look at readmissions to my center versus readmissions to all centers.

CO-CHAIR TRAVIS: Thank you.

CO-CHAIR BULGER: That's an absolute 3 percent or a relative 3 percent?

MS. MORSE: An absolute 3 percent.

CO-CHAIR BULGER: Any other questions on reliability before we vote? Okay, and then this -- like the last bunch we've done, moderate is the best you could do.

Sorry, Keith.

MEMBER LIND: So that -- I don't think that 3 percent was in the documentation. But I think that speaks, to me at least, that says a

lot about how much of this -- these numbers are going to change if you add in the other facility readmissions.

I mean I was imagining from your comments, Karen, that it could be, I don't know, 20 percent.

CO-CHAIR BULGER: Yes the only question I would have on the 3 percent if we want to go down that road is is that consistent across hospitals or is that the average, and some hospitals it's 20 percent and some hospitals it's, you know --

MEMBER JOYNT: And 3 absolute percent in readmission rate --

CO-CHAIR BULGER: That's a big deal.

MEMBER JOYNT: -- on a base of 15 percent is huge. Three absolute percent in the number of people coming to your -- right? I didn't totally understand what the 3 percent was.

MEMBER LIND: So that's like 20 percent of the readmission rate, is that what you're saying?

MS. JAGELS: Denise, can you answer that?

MS. MORSE: Yes, so this is for one time period and so we're looking at the all cause for -- honestly for my center the rate is 29 percent back to my center and it is 32 to all centers. So when looking at the all cause, we look at very large numbers for the cancer population because of the nature of cancer.

MEMBER LIND: And can you shed any light on the rate for other institutions than your own? Because that, obviously, is a concern.

MS. JAGELS: In Seattle it looks similar to the City of Hope numbers.

Susan, have you seen other absolute or modified rates for others?

DR. WHITE: No, I have not looked at that.

CO-CHAIR BULGER: You know, I asked too because I know just looking in our own -- in our system, and again, it's not cancer, it's all of them, but in the seven hospitals -- and we

track both because we track one with Premiere and we get the all cause from CMS, you know some hospitals it's 2 percent absolute difference, and again, it's on a 15, 16 percent Medicare. And some hospitals it's 5 or 6 percent.

But it depends. A lot of it is market share based and you know, some of them -- everybody comes back to the same hospital and others there's a hospital right down the street and wherever EMS happens to take them when they're really sick, that's where they go. So it depends on the -- which is why this is so variable which is -- to me, gives me a lot of concerns about both reliability and validity.

CO-CHAIR TRAVIS: Can I ask just one other clarifying question? Do you continuously get that data or was that just a one time that your hospitals got the dry run from CMS? Or do you get it every quarter, annually?

MS. JAGELS: Denise?

MS. MORSE: Yes. We do get that -- I believe it's annually that we receive that

report. We did not for a while and then we started receiving it again, so I think there's some variability in terms of whether we are included or excluded from their dry run set.

CO-CHAIR TRAVIS: So there is some information available on that, I guess, because the way that it was described was that there's not any information on where your patients go.

I know it's set up differently -- I mean there's a different construct, but it does seem like there's some information that gives you a feel for if they're going other than to your hospital.

MS. JAGELS: You're correct. A more nuanced answer on my part would be in measure development, we were using the UHC hospital claims data. The distinctions between that data set and what CMS provides us for our Medicare patients, we try not to mix them up --

CO-CHAIR TRAVIS: And I do appreciate that. Thank you. I understood a methodological kind of difference, but there's at least some

information that's out there.

MS. JAGELS: Absolutely true. Thank you.

CO-CHAIR TRAVIS: Thank you.

CO-CHAIR BULGER: Okay, any other questions or comments on this? Karen, go ahead.

Yes and I think this will be -- this is obviously, as we said, that this is -- if this is low, we stop. So I mean, as much discussion as we want to have.

MEMBER JOYNT: Yes, this may or may not be germane to the current thing, but I was just trying to find any other sources of data.

So in the Massachusetts hospitals, the range of readmissions to other hospitals by region, so metro Boston, 43 percent are to another region, and the places where people don't travel as far to get to care, it's between 7 and 10 percent are to another hospital.

So if your hospitals are similar types, you may have a narrow range. If they're not, you may have a huge range. It just seems to

me like this is a testable issue. Not easy to test, but either using Medicare data or an all-payer claims data or a sub-sample or something.

I don't have a feel for how big of an issue we're talking about. If your hospitals are very uniform, maybe the difference between them or the calculation of performance with or without is the same, in which case this is a non-issue.

MS. JAGELS: That's why today has been a huge learning experience. Obviously we didn't set out on a quest to avail ourselves to large sum ResDAC-type data. The question we were asking ourselves is for our hospitals, for our patients, in our claims data, could we successfully identify opportunities to better manage these patients outside the hospital?

So you're right, it -- once again, it's a grand assessment of where the patients are being readmitted and why. To other hospitals, this measure wouldn't be sensitive enough to detect that.

But we do believe for the purposes

that we set out to establish, could we, from a quality improvement perspective, find our patients, identify them, understand the frequency with which they're entering and exiting the hospital, for foreseeable and avoidable reasons? We think we did a decent job.

MEMBER JOYNT: Is the measure used to compare hospitals or only for individual hospitals to do their own quality improvement activities?

MS. JAGELS: So we are sharing our data across settings, but currently only using it for internal quality improvement. City of Hope, Seattle, and the James are robustly reporting by service line and by provider so that once again at the clinical decision making level, we can improve our performance.

But it's not a benchmark, nor are there targets, nor are we asking CMS to set those for us.

CO-CHAIR BULGER: Okay. Leslie.

MEMBER LESLIE HALL: So I think this

is a conundrum we'll have at any time we get presented a very specialized group of patients with a strong desire to use quality to improve your organization. I think it's fair to say that this is a learning effort as much as anything else. And I really support that idea. This is a group of patients that have opportunities to prevent readmissions and to participate and engage in those discussions, decisions, and they're very active patients.

So I think that we should consider this a learning effort and an eye-opening effort and maybe request, as a result, if this measure goes through, reporting back in a more frequent cadence to see if that measure has -- either had the unintended consequence that have been pointed out here, or actually been beneficial to quality improvement efforts. And that way that might mitigate some of the ongoing concern about both the narrowness of this approach and the unintended consequences of the other facilities that may be impacted.

CO-CHAIR BULGER: Taroon, you want to --

MR. AMIN: So on this issue of use, I think we need to clarify a little bit in terms of what the role of this committee is in terms of the endorsement process.

The endorsement process is intended to identify national performance standards for the purposes of quality improvement and accountability applications which would include public reporting. So we will need to address this question again if we get use and usability in terms of, you know, making sure that this measure is capable of doing both of those things.

And I'm not suggesting that it is or isn't. That's part of the committee's deliberations and it's not necessarily under the domain of reliability, but I understand there's concerns among the committee that they want to get all the concerns out before we vote on reliability. I just needed to clarify that specific point because it's important to

understand what the committee is voting on.

CO-CHAIR BULGER: Okay. Paul?

MEMBER HEIDENREICH: That was basically my point. I think there are definitely measures that are very important for quality improvement, but they're not yet at the level for public reporting and so, you know, we don't -- there's many measures we would potentially not even consider that are very important for communities and hospitals to use.

CO-CHAIR BULGER: Okay. Keith.

MEMBER LIND: Maybe the developers can clarify this, but my understanding is that these hospitals tend to specialize in different conditions. So to some extent the comparability, you wouldn't expect their readmission rates to be identical and the comparisons might or might not be valid. There's some overlap, as I understand it, but my understanding is that there is also -- they're reputationally specialized in different areas.

MS. JAGELS: I would agree, and in a

moment I'll let Denise speak directly to that. For instance, in Seattle and in Los Angeles, we attract many patients for the purpose of treatment for leukemia and lymphoma and bone marrow transplant. Naturally, the treatment trajectory for those particular modalities is in many cases highly intense, so we see patterns of patients entering and exiting the hospital that look different for patients undergoing colorectal or breast cancer therapy.

Denise?

MS. MORSE: Yes, that is absolutely correct. Our two sites tend to be the highest overall percentage of hematologic cases compared to other sites. And some sites within the PCHQR program do not treat or perform bone marrow transplants, as well, which is a huge differentiator. So even -- correct, within our group there are differences.

DR. WHITE: This is Susan, just for a second. We did make some effort to try to adjust for that in looking at -- in our risk adjustment

methodology, looking at solid tumor surgery and presence or absence of BMT -- bone marrow transplant status.

We at the James are also a pretty heavy hem-onc hospital and we have some predictive modeling that we do internally for readmissions and have found that that's a pretty good proxy for that mix. It's better than traditional comorbidity sometimes in predicting our readmission rates.

So not that it's a perfect model, but we did try to address some of that in our risk adjustments.

CO-CHAIR BULGER: Karen? All right. Any other questions, comments?

All right, so we're going to vote. Remember 1 is moderate. The highest you can get is moderate.

MS. HERRING: Voting is now open for reliability for Measure 2884. One, moderate, 2 low, 3 insufficient.

(Voting.)

MS. HERRING: The results are 5 moderate, 13 low, 1 insufficient. Just 26 percent moderate, 68 percent low, 5 percent insufficient. This measure does not pass and we will conclude voting on this measure.

MR. AMIN: So before we move on, I think it might be helpful just to summarize a little bit of the feedback for the developers because although we had to stop discussion at the point of reliability, the conversation spanned just reliability. So I think one of the -- just to -- I'll do a very abbreviated version of what I heard.

Related to the reliability testing, I think was what triggered some of the questions related to the .8 in terms of inter-rater reliability conversations about the underlying data element.

There was validity conversations related to readmissions to other facilities, and then I think there was probably some conversation around use and usability, which again related to

the validity question.

If there are any other sort of high level feedback elements, I'm sure the measure developers would welcome that.

CO-CHAIR BULGER: Yes, I mean the comment I would make is that, you know you made the comment you don't know what you don't know. And I think some of that you need to know before it would go through under the endorsement process.

So some of the stuff I think you can say, you know, you don't know what you don't know, but I think some of that it would be very beneficial. So, for example, this issue around what are the -- what is the difference between same hospital and even if you did a sample or you found some way to -- across your hospitals to understand what the spread was, knowing that information I think would be very helpful to the committee going forward.

And it doesn't necessarily -- you could know what the spread was and that spread

turns out to be equivalent, as Karen said, which would give you that, that you could absolutely say that was the case. But I think that would be very helpful to know.

Bruce?

MEMBER BRUCE HALL: Taroon, I think there was one other comment early on which might have got lost and that was that the title itself doesn't really reflect the fact that this is actually urgent emergent readmission and using the word unplanned actually I think draws too many parallels to the CMS unplanned algorithms which are not involved here.

I would, at the same time, suggest or ask whether this committee is permitted to sort of give this developer group an endorsement of some kind saying we think this is incredibly important work and we'd love to express to your constituents that the work should continue, right? And that might even help them get their work done. It might help them with local support and everything. So I just put that out on the

table.

MR. AMIN: We would certainly reflect that in the feedback to the developers in our materials. There's obviously a lot of great work that was done here by the developers and it's moving us significantly further in terms of measuring this important outcome for a population -- for this important care population.

So we'll definitely reflect that in the materials.

CO-CHAIR BULGER: Wes.

MEMBER FIELDS: Yes, I just want to support what Bruce suggested, whether it's official or it's just friendly. I actually gave you a moderate vote, but I think the discomfort of people that didn't centers on the fact that most of the measures we look at are more universal, all community hospitals, all ACOs, et cetera.

And I think what we all recognize is that you're essentially providing quaternary types of services that are regional in nature so

it's fundamentally different. But I think what we'd all probably like to see more is that even if you're quaternary in structure admission, you still interact in a number of ways not just with your patients, but with other providers.

And I think knowing what your readmission rate is to all facilities is worth knowing, even if it's not a truly significant number. Because I think where this quality movement is going overall is that we expect to have more and more special -- highly specialized procedures be regionalized, but we're getting to a place where the flow of patients back and forth for those highly specialized services is getting to be a pretty big deal.

So knowing when they bounce and where they bounce is important even -- and it's not because we disbelieve what your data shows, but in order to understand the best fit and the best outcomes in terms of authorizing readmissions, I think you really need to know what that interaction is with admissions away from your own

quaternary center really is.

CO-CHAIR BULGER: Okay, any other comments from the committee?

So thank you. I think we all just said -- I think you heard that the committee is very supportive of what you're doing and hopefully that continues. And I think we would be thrilled to see it come back at a future time.

MS. JAGELS: Well, I'd like to express my gratitude. Believe me, as I sat through the morning and had an increasingly urgent sense of unsettlement I thought okay, I can feel it coming. So I think your advice is very genuinely offered and I really appreciate the opportunity to even bring a measure of this stature before you. So thank you for your time.

CO-CHAIR BULGER: Thank you.

All right, so the next one is 0171 which is Acute Care Hospitalization During the First 60 Days of Home Health, it's a CMS measure. And the discussants are Helen, Paulette and Pam, and we'll start with the developers. Thanks for

coming.

MS. KEANE: So first of all, thank you for allowing us to present these measures and to answer any questions that you may have about these two measures. Again, it's 0171, Acute Care Hospitalization During the First 60 Days of Home Health, as well as Measure 0173, Emergency Department Use Without Hospitalization During the First 60 Days of Home Health.

Helping me present today, I have my colleagues. I have Dr. Jennifer Riggs, a nurse researcher at Abt Associates. I have Dr. Stephen McKean, who is our analytical lead for all home health claims based measures. And I also believe I have on the phone our measure steward, two representatives from CMS. And also I'll just offer Jennifer and I have both seen patients in home health as nurses in the past.

So both measures are currently endorsed and they're also outcome measure with the data source of administrative claims. These measures are for home health. Both measures are

currently used in the accountability reporting programs for CMS. For publicly reporting, these measures are reported on Home Health Compare. Measure 0171, what I'll refer to as ACH Measure, is used in the Quality of Care Patient Star Ratings program.

The home health acute hospitalization, or ACH Measure, and the Emergency Department Use Without Hospitalization Measure, or ED, are harmonized with the rehospitalization measures which are NQF numbers 2505 and 2380. And with CMS's hospital-wide, all cause, unplanned readmission measure, which is NQF 1789, and the definition of unplanned hospitalizations.

They do differ from other post-acute care hospital readmission measures, however, in the definition of eligible post-acute stays and the risk adjustment approach and by measuring emergency department use as an outcome.

The differences arise due to the unique nature of home health care as a post-acute setting. The ACH and ED Use Measures were

initially developed and later leveraged to construct the rehospitalization measures by further restricting the ACH and ED Use Measures' eligible population by requiring prior proximal in-patient hospital stay within five days from the start of home health.

Finally, both pairs of measures are risk adjusted using patient level predicted probabilities calculated from multinomial logistic regression.

Risk factors that are accounted for in both pairs of measures include demographics and health status as measured by both the CMS hierarchical condition categories, HCCs, found on claims in the previous six months, the rehospitalization measures leverage the prior proximal in-patient hospital claim to obtain the patient's diagnosis related group or DRG, and also risk adjust for the activities of daily living fields on the OASIS, or Outcomes and Assessment Information Set assessment of the initial home health stay.

The risk-adjusted rates for the ACH and ED use measures are publicly reported, as I previously stated. However, due to a large number of relatively small home health agencies treating previously hospitalized patients, the measure developer determined that reporting home health agencies' risk-adjusted rates could lead to misleading conclusions since small home health agencies' risk-adjusted rates tend to be unstable.

Therefore, the risk-adjusted rates for the home health rehospitalization measures are publicly reported as categorizations, for instance, better than expected, same as expected, worse than expected.

While the acute-care hospitalization and Emergency Department Use Without Hospitalization Measures differ from other post-acute care measures in some regards, these differences arise from the unique nature of home care as well as a desire for harmonization across all of our home health quality measures.

The only thing I also want to call out, as there were no major changes in these two measures since they were previously endorsed in 2012, we have had two minor changes that were made to the measures. First, the title of the measures were changed to improve the clarity. Previously, they just said acute care hospitalization or emergency department utilization. The first 60 days of home health has been added to clarify that, which I'm quite happy we did after the previous presentation.

Secondly, there's been a recalibration of the risk adjustment model coefficients using data from January 1st, 2013 through to December 31st, 2013.

So I don't know if either of you want to add anything.

MEMBER CHEN: So as the developers mentioned, this is a maintenance discussion and this measure has already been in use in the community for public reporting.

In terms of evidence, the evidence

that's mostly provided in this measure is really around readmissions. There's a mention of some more recent work at the -- from QIOs about community-based interventions that also prevent acute care hospital utilization. To be honest, I couldn't get into that study, but I did look at the Jane Brock, Joanne Lynn JAMA paper from 2013 and they did show an ability for home care agencies and communities to actually intervene on all admissions, not just readmissions. So the evidence does exist out there for this.

CO-CHAIR BULGER: Vote on the evidence.

MS. HERRING: Voting is now open for evidence for Measure 0171. Your choices are 1 yes, 2 no.

(Voting.)

MS. HERRING: The results are 18 yes, 0 no. So 100 percent yes.

CO-CHAIR BULGER: Gap?

MEMBER CHEN: In terms of the performance gap, the developers reported new data from 2011 to 2014 regarding performance rates.

For 2014, the interquartile range was 12.7 to 16.8 percent. And also in the validation set the difference between the 10th and 98th percentile was 11.3 to 22.9 for those agencies with at least 20 stays. So there is a performance gap.

In terms of SDS, they did report on a conceptual rationale. They did look at some of the measures and as discussed in other measures previously, they recommended not including SDS in the model.

CO-CHAIR BULGER: Great. Thank you. Pam?

MEMBER ROBERTS: I think the only thing to add is that they did note variations across facilities.

CO-CHAIR BULGER: Okay. Any thoughts from the committee? All right, we'll vote on the performance gap.

MS. HERRING: Voting is now open for performance gap. For measure 0171, your choices are 1, high; 2, moderate; 3, low; 4, insufficient.

(Voting.)

MS. HERRING: The results are 3, high; 15, moderate; 0, low; 0, insufficient; so 17 percent, high; 83 percent, moderate.

CO-CHAIR BULGER: Okay, reliability. Pam, do you want to --

MEMBER ROBERTS: They did do reliability tests, as I mentioned, using the beta-binomial. They did it at the patient level and the reliability score was greater than .871 and at least 50 percent of the agencies had a reliability score of .77. So they had some good reliability.

CO-CHAIR BULGER: Okay, Helen? All right. Anything from the committee on reliability? Seeing none, let's vote.

MS. HERRING: Voting is now open for reliability for measure 0171. Your choices are 1, high; 2, moderate; 3, low; 4, insufficient.

(Voting.)

MS. HERRING: The results are 2, high; 16, moderate, 0, low; 0, insufficient. So 11

percent, high; 89 percent, moderate.

CO-CHAIR BULGER: Excellent. Validity, or excuse me, yes, validity. Helen?

MEMBER CHEN: Data element validity was not tested because CMS audits the sample for accuracy of claims and the claims are very accurate. They did perform a random split sample of the agencies with at least 20 stays with 80 percent of the facilities in the development group and 20 percent in the verification group. And the c-statistic was actually 0.693 in both samples and the cross-validation at the 10th and the 90th percentiles for predicted probabilities was 8 and 31 percent in both.

CO-CHAIR BULGER: Anybody from the committee with comments? Karen?

MEMBER JOYNT: I just have a question. So I know you guys have two measures and they're sort of related to each other. It looked like there's a beta in this measure for whether or not you've had an ER visit. Is that right or did I misread that? I'm trying to understand how the

two measures work together that the revisit versus the rehospitalization. Did I understand that that event distributes here?

DR. MCKEAN: Right, so there's a multinomial logit model, it's both measures, basically, for the two different outcomes. So you would have different parameters estimates predicting the -- you would have two sets of parameter estimates. One set of parameter estimates would predict whether or not you have the ED use without the hospitalization and then the other set of parameter estimates would predict if you had the acute hospitalization. But it's one multinomial logit model predicting the two outcomes.

MEMBER JOYNT: So whether or not you have a -- is there actually a term in the hospitalization model that says whether or not you had an ED visit? Or did I just misunderstand what that said?

DR. MCKEAN: So the parameter estimate in the risk-adjustment model would have.

MEMBER JOYNT: Okay.

DR. MCKEAN: There might be a parameter estimate in the risk-adjustment model, but I would say if you had previous ER visits, I could pull up --

(Simultaneous speaking.)

MEMBER JOYNT: Oh, I see. It's not about -- it's not because the model is linked, that it's actually built that way. It's whether or not you had a prior ED visit.

DR. MCKEAN: Right.

MEMBER JOYNT: Is the thing that informs this model. Okay. I was thinking about this is a bit of a usability question, also a validity question, but it's a little hard to get your head around how to think about -- is it good if you have low readmission rate, but a high ED use rate and --

DR. MCKEAN: The goal would be to have a low for both of them and that could be possible. And that is possible.

MEMBER CHEN: Just as a point of

clarification, this all acute care hospitalization and not just readmissions? Yes.

CO-CHAIR BULGER: Okay.

MEMBER AUGER: Sorry to follow up on that, but it is one model for both measures, is that correct?

DR. MCKEAN: Right, that is correct.

CO-CHAIR BULGER: Bruce.

MEMBER BRUCE HALL: I was trying to get my head around perhaps a related issue and I don't know whether necessarily it's validity or use, but as an unintended consequence, if a facility just uses home health for everything, they're going to look great, right? Okay.

CO-CHAIR BULGER: Other questions? Karen.

MEMBER JOYNT: This is another related issue that I think relates to both validity and maybe unintended consequences which is you also have to be accepted into home health, so whether or not home health is available and whether or not you're accepted and whether or not you sort

of make traces between where you're sent. It just makes me a little bit nervous that the selection into an exposure is very different here than it is in many of the other settings we look at. I don't know that I actually have a prior and what to do about it or how it might influence things, but the selection feels important to me in judging how well you can really tell patients that are in home health agency versus another given that it's a little bit like that sort of selection bias of treatment assignment that we always worry about in observational data.

MS. KEANE: So I will offer that there is patient choice to select which sites they go to, so if a patient is making a choice to go to home health, they're making a choice to go to home health. Beyond that, I think this is a home health measure.

CO-CHAIR BULGER: Carol.

MEMBER RAPHAEL: I have to respond to that. I mean yes, there is choice in the regulations, but in reality, one of the major

issues is that many people land in nursing homes because it's a Friday afternoon and there's an available bed and the hospital wants to discharge someone who's medically ready to discharge. And before they know what hit them, they're in the nursing home. So I really think that is an issue that we should be cognizant of. And people who land in home care often don't really have a choice. I mean yes, they're given a list, but they ask what is your recommendation? And they often will go with the recommendation of whoever the hospital kind of has preferred provider relationships with. I don't know how that plays out in terms of selection, but that's how it operates in the world that we inhabit for better or for worse.

CO-CHAIR BULGER: Other questions or comments? I just would say that's evolving a lot, too, with the narrow networks, you know, the previous ACO discussion and as networks narrow, the choice -- again, there's still choice, but you're given -- the list has narrowed certainly.

Keith?

MEMBER LIND: So the agencies also have a choice.

MEMBER RAPHAEL: That's what I was talking about.

MEMBER LIND: That's what I thought you were talking about. So it's not just the patient. The agencies have a lot more flexibility about whether or not to accept a patient than a hospital does when they're deciding to admit an unplanned admission. There's no EMTALA for health agencies. So I don't know how you deal with that. You can adjust for clinical and patient characteristics, but there's an inherent selection bias in that measure. But I mean it's already approved, but I wanted to mention that.

CO-CHAIR BULGER: Any other comments? Pam.

MEMBER ROBERTS: I think you're starting to see though in some of the markets, at least in urban markets when you have quaternary

hospitals, they are really working with home health agencies that can handle those level of patients. And you're starting to see much more coordination of care starting to happen. So I think there could be some good byproducts of this that are starting to happen I guess over time. We'll see more, but at least I can say for urban markets you start to see that.

CO-CHAIR BULGER: Anything else? Okay.

MS. HERRING: Voting is now open for validity for measure 0174. Your choices are 1, high; 2, moderate; 3, low; 4, insufficient.

(Voting.)

MS. HERRING: The results are 1 high; 17 moderate, 1 low, zero insufficient. So 5 percent high, 89 percent moderate, 5 percent low.

CO-CHAIR BULGER: Okay. Feasibility. Helen?

MEMBER CHEN: It's basically claims data. That's fine. Not much to say.

CO-CHAIR BULGER: Any other comments?

Okay, we'll vote on feasibility.

MS. HERRING: Voting is now open for feasibility for measure 0171. Your choices are 1, high; 2, moderate; 3, low; 4, insufficient.

(Voting.)

MS. SHAHAB: Tom, can you submit your vote, please?

MEMBER SMITH: I texted you a 1.

(Laughter.)

MS. HERRING: Thank you. And the results are 16 high; 3 moderate; 0 low; 0 insufficient. So it's 84 percent high; 16 percent moderate.

CO-CHAIR BULGER: Use, Pam. Do you have any comments on use?

MEMBER ROBERTS: I mean they're starting to show some risk-adjusted performance. It's very slight at the agency level. And it's been stable across the population level.

CO-CHAIR BULGER: Helen.

MEMBER CHEN: It's being used.

CO-CHAIR BULGER: Any comments from

the group? All right. Oh, Cristie.

CO-CHAIR TRAVIS: Any expectation that we'd see more change than perhaps we're seeing with this measure? Are there any particular relevant issues that we need to see? If I'm reading this correctly it's stayed about the same, even though it is being used. I guess I was hoping that we would see an improvement over time. So just any thoughts you have about that, I would appreciate it.

DR. RIGGS: I think one of the things that's holding us back a little bit is that there's not a great deal of research that's home health care specific. And so we're not necessarily -- we don't necessarily have the evidence that we need to take us to the next level. However, as we speak, there's an awful lot of activity going on really trying to identify what best practices are that really will move the needle in these areas. And so I think over the next five years, we may, in fact, start to see an additional step forward in terms of

reducing these rates. I agree. I see exactly what you see. We're kind of plateaued here, but I think we're just moving a little bit slowly because of the nature of the care setting.

CO-CHAIR TRAVIS: Thank you.

CO-CHAIR BULGER: Okay, any comments? Paula.

MEMBER MINTON-FOLZ: Do you see much of your readmissions from patients who really did not want to go into a skilled nursing and now need to and that three-day rule? Is that part of that you have to have -- or is that just Washington where it's a three-day rule before you -- oh, okay.

Well, if a patient needs to go into a SNF from home, they need to have a prior three days of acute care hospitalization. Do you see the impact of that affecting your readmission rates? Do you see patients often going from home to hospital to SNF? Or are they often going home, hospital, back home? Is that clear?

MEMBER FIELDS: Paula, I think I can

answer the question for you even though I don't know their data. But I think you're talking about a different subset because the ones you're worried about and it's a legitimate concern is the patient that doesn't meet criteria for admission for three days of in-patient service is not well enough or stable enough to be independent at home. Therefore, it gets referred to home health services that prove to be inadequate. But that won't show up in their data because they don't have that index admission.

CO-CHAIR BULGER: Other questions? Okay, we'll vote on usability, use.

MS. HERRING: Voting is now open for usability and use for measure 0101. Your choices are 1, high; 2, moderate; 3, low; 4, insufficient.

(Voting.)

MS. HERRING: The results are 3 high, 16 moderate, 0 low, 0 insufficient. So 16 percent high, 84 percent moderate.

CO-CHAIR BULGER: Okay, that brings us

to the final question, the overall suitability. Any further -- Pam, Helen, anybody from the committee? Okay.

MS. HERRING: Voting is now open for overall suitability for endorsement for measure 0171. Your choices are 1, yes; 2, no.

(Voting.)

MS. HERRING: The results are 19 yes, 0 no. So 100 percent yes. Thank you.

CO-CHAIR TRAVIS: Okay, well, we're ahead of schedule which is great and we're going to come to our last measure now which is 0173, Emergency Department Use Without Hospitalization During the First 60 Days of Home Health.

Are there any additional comments that you all wanted to make? Okay. The developers have already made their comments earlier. We've got Wes, Carol, and John as our lead discussants. So Carol, would you like to go first?

MEMBER RAPHAEL: I just wanted to make a kind of overall comment which is I know this has already been endorsed in the past and it's

being used for public reporting. But I did have to say and maybe this is in line with a point that was made earlier, I thought the research base for this was very thin. There was one study that really kind of was the buttressing study. And there's very little evidence of what can make a difference in regard to really preventing unnecessary emergency department visits.

The only things that I have seen and it's still very early stage references to telehealth, but I don't think we know enough yet. Access to primary care, which is really tough because you get patients who don't have a primary care physician, a primary relationship. And then even if they do and they call, the primary care physician is not available. And then all of the things that we thought made a difference like medication, reconciliation, a lot of kind of front-loading, follow-up visits, patient education, and activation and all of that good stuff, falls prevention, makes no difference whatsoever.

So when you put this all together, I have to say I felt like I was standing on kind of sand rather than stone. And so that really did concern me in terms of thinking about how do get improvement here and where are we headed with this? But I don't know, Wes, what your reaction --

MEMBER FIELDS: Yes, I want to disclose tow conflicts. One is my group now offers a telemedicine service and the other is that since this is a maintenance measure that I probably can't kill, I think I'd rather catch my plane. But that being said, that being said, you know, the developers have actually given the best reason for this to be maintained. And it essentially functions as a tracking measure for the interaction between condition at hospital discharge, condition at referral to home health services, and the likelihood that they'll bounce back to the emergency department. So we all want to prevent preventable readmissions, and I think that since this is kind of an area that's more

dynamic in the marketplace, as you say, rather than pure research, I think we need to continue to follow this, but that's probably it's only reason to exist, including your statistical model.

But I do think we spent a fair amount of time yesterday talking about Medicare patients' condition on discharge from acute care being significant, a significant issue in terms of their ability to provide self-care or to be stable for community-based care with home health. And I think that's the only reason that this is worth continuing to follow because there is no science here. And even the one reference that you have is really about frequent visitors to the emergency department which is a really -- that's a different population.

But I do think we should support it, because I think we need to do our best to make sure that people being discharged from acute care are in sustainable health statuses in the community.

**CO-CHAIR TRAVIS: John, anything else?**

**CO-CHAIR BULGER: I feel the same way, but I won't belabor -- I think you're both exactly right.**

**CO-CHAIR TRAVIS: This is a maintenance measure and the developers have indicated that the underlying evidence has not changed, but I want to get a read from the committee as to whether we want to kind of vote on the evidence criterion? You know, I'm just kind of hearing both sides of the issue here and I just want to be sure I'm clear on what our next steps should be.**

**MEMBER FIELDS: Well, I think you should consider that the sponsor of the measure is CMS and that this is a moving target. I think we should probably move on from here and to support the measure's renewal.**

**CO-CHAIR TRAVIS: Pam.**

**MEMBER ROBERTS: I think that we're going to need to watch this measure over time and we should support it now because especially with**

all the bundling payments and the changing payment models, this could become a very important measure if people start using the ED or if we can if we can really keep them out.

CO-CHAIR TRAVIS: Okay, well, not seeing anybody say we should vote on it, oh, I'm sorry. Kathy, I didn't see your card.

MEMBER AUGER: The only other evidence, sort of type statement I would make is that it seems like sometimes home health care referrals to ED is not seen as a bad thing. It's actually seen as a good thing that the home health care agency or home health care nurse was in the home and recognized a problem early and was able to get them to the appropriate level of care. So that's just the other part of me. And it ties up, perhaps, this multinomial model that they're going to the ED, but not getting readmitted, so that may not be a bad thing. That's the only --

CO-CHAIR BULGER: And I think it's good that you track it, too, for the reasons that

were mentioned because I think you're going to see -- if you think of a curve of -- as opposed to acute space is one that everybody is looking at because if you look at where the variation is in Medicare it's in that post-acute space after hospitalization. If you want to save money and you're an ACO, after not getting them in the hospital in the first place, where they go is important.

So this kind of curve with -- after hospitalization with say LTAC on the far left being the most expensive, and then in-patient rehab and then SNF and then home health, people are going to start trying to shift that curve to the right. So you're going to see sicker and sicker patients, if you will, going to home health.

So I think this notion is well said of having this to track because you'll be able to pick up these things and Pam mentioned it, too, and so did Carol. I think it's important to have it as a tracking device to be able to -- it's

almost maybe a canary in a coal mine to see if things are starting to go wrong with that whole process.

CO-CHAIR TRAVIS: Well, I think based on this discussion and the advice of Taroon, it probably would be best for us to go on the record of voting relative to evidence. And this a must-pass criterion.

MS. HERRING: Voting is now open for evidence for measure 0173. Your choices are 1, yes; 2, no.

(Voting.)

MS. HERRING: So the results are 17 yes, 1 no, so 94 percent yes, 6 percent no.

CO-CHAIR TRAVIS: Okay, thank you. We'll go to performance gap, and Carol, you want to start us out there? Can you put on your microphone? Thank you.

MEMBER RAPHAEL: I don't think there was anything that was noteworthy. We're not making progress. I think that was said before and we have a gap to close here. I don't think

there's anything more to add.

MEMBER FIELDS: No, that's actually a little bit alarming. The results look the same and I think we'd all like to see them get better. And I think we all support CMS' intent with this measure. But the reality is it's not happening yet.

CO-CHAIR TRAVIS: John, anything to add?

CO-CHAIR BULGER: I don't other than to say is there some idea as to why that's the case? Is there anything in the -- any thoughts?

MEMBER RAPHAEL: The only point that was made and I don't know if you can comment on it, we said anecdotally, there seems to be evidence that -- and this is a line drawn with where you were headed -- that the patients now in home health are older, more women who are living alone tend to be sicker. That would be the case --

CO-CHAIR BULGER: That was one of my -- you made it. It is getting better. But the

patient population is getting sicker and those two things are equaling each other out so that you end up staying the same.

DR. BURSTIN: But that's the data you should have.

MEMBER RAPHAEL: Helen, it was mentioned it was anecdotal evidence.

DR. BURSTIN: I am just saying that you would think that they would have the data to, in fact, look to whether the population has shifted in terms of age, number of comorbidities, etcetera. If they have the whole system.

DR. MCKEAN: Just pulling up, we have numbers that show the distribution of the age groups by year, so from 2011 to 2014. And it does look like the biggest group which is the 75-to 84-year-olds, that stays relatively constant, around 35 percent over time. So it doesn't look like the distribution of ages is dramatically changing over time. So that might not necessarily be what's driving this.

MEMBER FIELDS: Can you tell us if --

DR. MCKEAN: There could be other comorbidity and distribution --

MEMBER FIELDS: So there's no data for comorbidity or MCC proxies.

DR. MCKEAN: We could look at that over time.

MEMBER FIELDS: Well, I think that's kind of what we're suggesting.

DR. MCKEAN: All I have right now in front of me is distribution by age, sex, race.

CO-CHAIR TRAVIS: Any other conversation on performance gap? Time to vote.

MS. HERRING: Voting is now open on performance gap for measure 0173. Your choices are 1, high; 2, moderate; 3, low; 4, insufficient.

(Voting.)

MS. HERRING: The results are 4 high, 13 moderate, 0 low, 0 insufficient. So 24 percent high, 76 percent moderate.

CO-CHAIR TRAVIS: Okay, now we'll move to reliability. Wes, any comments on

reliability? Carol, John? Does anybody on the committee have any questions or comments on the reliability testing? Okay, we can go to vote.

MS. HERRING: Voting is now open on reliability for measure 0173. You can choose 1, high; 2, moderate; 3, low; or 4, insufficient.

(Voting.)

MS. HERRING: The results are 0 high, 17 moderate, 0 low, 0 insufficient; so 100 percent moderate.

CO-CHAIR TRAVIS: Validity. Wes, anything?

MEMBER FIELDS: I think we've beat up validity pretty well. But I really do think that this is a companion to the readmission piece and as long as your model doesn't change, I think we need to continue to track this. But I'm hoping to hear that our real concern is for the population you serve and whether it's changing or becoming more unstable at the time of referral to home health. So if there's ways to enhance the model, I mean, what do they call it? The

continuity of care document, the electronic document following the patient? Yes, that should be a rich source of information for you to know what the health status of the patient is at the time of their discharge from the hospital. And it's pretty much of an electronic standard. So you probably have to reengineer that or rebuild it. But I think even if you can look at the parallel problems of the MCC population or rather high-risk, high-cost populations other than age, I think we'd all like to know what's happening in terms of the likelihood that they're referred to home health and the likelihood of them being referred back to the emergency department or even needing to be readmitted.

CO-CHAIR TRAVIS: Anything, John? Carol said no already. Okay, any other questions? Ready to vote.

MS. HERRING: Voting is now open for validity for measure 0173. Your choices are 1, high; 2, moderate; 3, low; 4, insufficient.

(Voting.)

MS. HERRING: Results are 1 high, 16 moderate, 0 low, 0 insufficient; so 6 percent high; 94 percent moderate.

CO-CHAIR TRAVIS: Okay, feasibility. Carol, any comments?

MEMBER RAPHAEL: I think it's feasible. It's based on claims.

CO-CHAIR TRAVIS: Wes is nodding his head. Okay, John, anything extra? Okay. We're ready -- I don't see any other cards, so we're ready to vote.

MS. HERRING: Voting is open for feasibility for measure 0173. Your choices are 1, high; 2, moderate; 3, low; 4, insufficient.

(Voting.)

MS. HERRING: I think we're just waiting on one more. Okay. We have 14 high, 2 moderate, 0 low, 0 insufficient; so 88 percent high, 13 percent moderate.

CO-CHAIR TRAVIS: Okay, use and usability. Carol, Wes, John, any comments you want to make?

MEMBER RAPHAEL: Well, I think as was said, it's already being used for home health compare.

CO-CHAIR TRAVIS: Okay, seeing no other discussion, we'll go to vote.

MS. HERRING: Voting is now open for usability and use on measure 0173. Your choices are 1, high; 2, moderate; 3, low; 4, insufficient.

(Voting.)

MS. HERRING: The results are 7 high, 9 moderate, 0 low, 0 insufficient; so 44 percent high, 56 percent moderate.

CO-CHAIR TRAVIS: Any final comments before we vote on overall suitability for endorsement? Okay. Thank you. We'll go to vote.

MS. HERRING: Voting is now open for the overall suitability for endorsement for measure 0173. Your choices are 1, yes; 2, no.

(Voting.)

MS. HERRING: The results are 16 yes, 0 no; so 100 percent yes. This concludes voting

for today. Thank you very much.

MR. AMIN: So just a quick thing before we move on from this measure. We did identify that it's related to 2505, the emergency department use without hospital readmission in the first 30 days of home health.

Just a quick reminder this committee reviewed that measure and this measure in terms of related and competing. The developers offered a rationale in terms of why both measures are needed and this committee agreed with that rationale as we review 2505. So it doesn't appear that we need to have an additional conversation about competing measures discussion unless anyone feels otherwise.

MEMBER FIELDS: Taroon, actually, remind me why we need both 2505 and this one? What is the rationale?

MR. AMIN: The developers contended that there are differences in justifying the two separate measures; patient admission to an emergency room without hospitalization during 60

days following the start of home health; and 2505 evaluates admission to emergency room within 30 days of starting home health for patients who are recently discharged from an in-patient setting. So I believe the idea was that 0173 assesses the efficacy of clinical care for all patients and so I just want to confirm that with the developers. Was that sufficient?

MEMBER FIELDS: Does 2505 not -- does 2505 not include an index referral after an in-patient stay? Is that the difference? One is referral from the community and one is a post-discharge thing? Is that the difference?

MEMBER RAPHAEL: From what I understand, that is the difference. The first one is 60 days post-discharge from an in-patient setting. The second is voted ED from the community and you are not --

MR. AMIN: Okay, thank you.

DR. LEVITT: This is Alan Levitt from CMS. The 2505 is ER use after hospital discharge. It's a readmission measure. It's 30

days and it harmonizes with the hospital-wide readmission measure. It includes only hospital discharged patients who go into home health.

The measure that we're talking about now is a hospitalization measure during the entire home health episode for all home health patients. I don't know the exact number, but probably 50 percent of the home health admissions are hospital -- come from the hospital. The other 50 percent come from the community. So it's a much larger group of patients that we're talking about. And it's over the home health episode of care which is that's where 60 days come from.

MR. AMIN: So the only other question I have for the committee in terms of relating and competing, we identified two measures during this morning discussion related to admission rates for heart failure and diabetes.

Paul, you had mentioned yesterday that there were other measures that you wanted to raise related to this related and competing

measure as it relates, I believe, to --

MS. SHAHAB: Between the access days and the readmission rates.

MR. AMIN: And I wanted to know if that was still a question, obviously not for right now, but if that needed to be raised for an additional conversation during our next conference call.

MEMBER HEIDENREICH: I think it should be touched on briefly in one of the calls.

MR. AMIN: Okay. Thank you.

CO-CHAIR TRAVIS: So we're going to open it up now for public comment. Is there any in the room? Seeing none in the room, Operator, will you see if there's any public comment on the phone, please?

OPERATOR: At this time, if you'd like to make a comment, please press star and the number 1. There are no public comments at this time.

CO-CHAIR TRAVIS: Thank you. Well, thank all of you for your time. Before we let

you go, however, we're going to have Zehra kind of walk us through what the next steps are.

MS. SHAHAB: Thank you, Cristie. So on the next slide you'll see a few dates. The first one is the post-meeting follow-up call which is coming up really quickly, June 21st and that's from 2 to 4 p.m. After this, we will -- after today's meeting, staff is going to be writing the draft report and we will have that draft report posted for public and member comment for 30 days from August 1st to August 30th. And then we will have a post-draft report call with you all, the committee, October 5th. And then the draft report will be posted for NQF member vote October 11th through 31st. And then we plan on going to the CSAC in November of this year and that's also when we will go for endorsement of the Board.

And finally, our appeals will be December 2nd through January 2nd, a 30-day appeals periods. And with that I want to thank all of the measure developers for their

incredible work. I know many of them are not here right now, but we want to thank you especially for all of the innovative work you've been doing with the SDS trial, responding to all of the committee feedback and the concern. And thank you for taking the time over these past two days for being with us and to answer all of the questions.

Second, I wanted to thank all of you committee members for such rich discussions and all of the work you all have done since there has been an extra load of work on your plates for this project and I can't thank you enough for making this project and all of our meetings together so enjoyable.

You're definitely my favorite committee hands down. I'm not sure if I'm allowed to say that, but I don't think I've ever said that before. This is the first time.

DR. BURSTIN: And you may have contributed to her decision to try to go to medical school, so there you go.

MS. SHAHAB: A very special thank you to all of our chairs, Cristie, John, and Bruce, and especially Bruce, who offered to help when John had a conflict on the first day and then I want to see if any of my staff or the chairs or any or all of you have anything to say. Taroon, Erin?

MR. AMIN: No, just thank you.

MS. O'ROURKE: Just to echo Zehra's thanks. She put it very eloquently. So thank you all and thank you to our developers and especially to our chairs.

DR. BURSTIN: Thanks. It's been a tough slog, but we couldn't imagine doing it with anybody else. So thank you.

CO-CHAIR TRAVIS: Or anybody else doing it. Thank you all very much.

MR. AMIN: Safe travels.

CO-CHAIR TRAVIS: This concludes the meeting.

(Whereupon, the above-entitled matter went off the record at 2:40 p.m.)

**Q**

**U**

C E R T I F I C A T E

This is to certify that the foregoing transcript

In the matter of: All-Cause Admissions and Readmissions Standing Committee

Before: NQF

Date: 06-09-16

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