NATIONAL QUALITY FORUM

+ + + + +

HEALTH AND WELL BEING
STEERING COMMITTEE MEETING

+ + + + +

WEDNESDAY
April 30, 2014

+ + + + +

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

PRESENT:

THOMAS MCINERNEY, MD, Co-Chair
SARAH SAMPSEL, MD, Co-Chair
CHISARA ASOMUGHA, Centers for Medicare & Medicaid Services
JOHN AUERBACH, Northeastern University
RON BIALEK, Public Health Foundation
JUAN EMILIO CARILLO, Weill Cornell Medical College, NYP
JANE CHIANG, American Diabetes Association
ERIC FRANCE, Kaiser Permanente
RENEE FRAZIER, Healthy Memphis Common Table
RON INGE, Delta Dental of WA
DAVID KROL, Robert Wood Johnson Foundation
MARGARET LUCK, Mary's Center for Maternal & Child Care
PATRICIA MCKANE, Michigan Department of Community Health
AMY MINNICH, Geisinger Health System
JACQUELINE MOLINE, North Shore Long Island Jewish Health System
MARCEL SALIVE, National Institute on Aging

KATIE SELLERS, Association of State and Territorial Health Officials
JASON SPANGLER, Amgen, Inc.
MIKE STOTO, Georgetown University
ROBERT VALDEZ, Robert Wood Johnson Foundation Center for Health Policy
ARJUN VENKATESH, Yale University School of Medicine

NQF STAFF:

HELEN BURSTIN, Senior Vice President, Performance Measurement
ADEELA KHAN, Project Manager, Performance Measurement
ASHLEY MORSELL
ELISA MUNTHALI, Managing Director
KAITLYNN ROBINSON-ECTOR

ALSO PRESENT:

MARY BARTON, NCQA
SEPHEEN BYRON, NCQA
ROBYN NISHIMI, Healthcare Quality Consultant
PAMELA OWENS, AHRQ
PATRICK ROMANO, UC Davis, AHRQ*
CAROL STOCKS, AHRQ*

* present by teleconference

A-G-E-N-D-A

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

9:07 a.m.

MS. MUNTHALI: Good morning again and welcome to the Health and Well Being Standing Committee meeting. This is day 2. And I will turn it over to our co-chair Sarah Sampsel.

But before we do that, Cathy, we just wanted to make sure that Patrick Romano from AHRQ is on the line.

OPERATOR: He has not joined yet.

MS. MUNTHALI: Okay. And that's fine.

DR. SAMPSEL: Well, good morning, everybody. And I guess another day in kind of gloomy, rainy D.C. But I'll speak on behalf of New Mexicans that it's really nice to see rain. That stuff falling from the sky.

Anyway, hope everybody had a great dinner and a nice evening last night.

And we managed to get a lot done

yesterday. However, today we do still have six measures. Five of those continuation of our PQI or prevention quality indicator discussion from AHRQ. And then an NCQA breast cancer screening measure before we go to some more general conversations about measure gaps and next steps for this committee.

But before we do that, just wanted to ask if there were any additional process questions, any reflections overnight that anybody wanted to share with the committee before we got started.

Hearing none and seemingly everybody is awake we will go ahead and get started. And our first measure this morning will be 0728, the asthma admission rate. In our PDF of all of the comments that's page 43.

And to start out I will ask Pam if she has any introductory comments or any reflections from yesterday's discussion.

MS. OWENS: There's not a lot more

to say than what was said yesterday. All of the measures that AHRQ is the steward for are in the same vein as yesterday that we talked about in terms of the overarching prevention quality indicator.

The measure you voted on yesterday was a pediatric measure. This is also a pediatric measure. The rest of the measures today are adult measures. So just making sure you understand we're changing age groups after this measure.

Again, the attribution is not assigned and it is not intended to be to a physician or to a hospital. It is a healthcare system broadly defined and that does include community factors.

And it's a way to drill down a little bit further and say maybe this is an area of emphasis or priority. I'm just reflecting on that because of the way the care setting statement probably from our paper or from what we submitted, it seems to

be much more specific and just highlighting ambulatory care, that's very broadly defined. So I'm just highlighting that piece for you.

All of it does use the healthcare cost and utilization project data. I gave you a little bit of description around there. So other than that I think we're ready.

DR. SAMPSEL: So, Amy, were you the lead on this one? No? Was this changed? Jacki and who? Okay, so Jacki, are you prepared to lead us through this?

MS. MOLINE: I had my name next to this one so I'm cool with this one.

(Laughter)

DR. SAMPSEL: All right, awesome. Okay, gotcha.

MS. MOLINE: So, you don't get the look of like a deer in the headlights like yesterday.

This is an update of an outcome

measure that is looking at admission rates for children age 2 through 17 for asthma. And it is a, again, this is looking to see whether improvement in the measure leads to less hospitalizations with the thought that it would be leading to better control and better management in an outpatient setting of an ambulatory-sensitive condition.

So, it was -- it's on a population basis, county or city. And it uses hospital data and administrative claims.

And I'm happy to go through the evidence if we'd like, or if you'd like me to do that.

Basically it's an avoidable admission looking at the rates which have been fairly stagnant. If you look at the rates through 2007 to 2011 the rates have basically not changed. There have been little blips. There is a definite difference between age grouping with the younger children being admitted more, and a

definite men -- boys are -- there's a gender disparity and there's also an income disparity. There's also a regional disparity which has not been well described or well -- it wasn't really addressed at all.

I mean, it's almost 3 to 1 from the Northeast to the West, whether it's environmental factors or other elements. I can't believe that asthma rates are so different. But that's what the data are.

And it's -- so that is basically -- there's definite room for improvement. There has not been much improvement since 2007 when you're looking at the data.

And I think this is a health priority. This is one of the leading causes of preventable admissions in children and a way of assessing whether better outpatient control can lead to decreased admissions.

And it is a marker in many ways of access to care and the overall ambulatory

health system. As well as what we discussed yesterday with the social determinants of health. Because it's such a multifactorial issue. And obviously some children will have greater severity of disease than others that could lead to it.

So, that is basically the introduction. I don't know, Margaret, if there's anything you wanted to add or anyone else?

MS. LUCK: Just to point out that during the committee call a lot of the issues we discussed yesterday about what's the utility of this measure, to what extent would changes in this measure reflect changes in the outpatient care system versus environmental changes. All of those issues were brought up in the committee call. But I think we talked about them.

MS. MOLINE: And the interesting thing about this measure for me was when you compare it to the gastroenteritis in

children is there was a tremendous drop-off about two years ago in the gastroenteritis admissions. And so the question was was that related either to the vaccination or was that related to differential coding because people were in a holding area in an ER. And so that they wouldn't actually be admitted.

This one there's no difference in their admission rates. So the severity of necessitating an actual admission has not changed and there is tremendous room for improvement.

MR. AUERBACH: So, I hope this doesn't feel like beating a dead horse, but I do think this is one where it really does make sense to say that the language I think in the proposal that refers to this is related to ambulatory care primarily is just wrong.

I think the data are really strong that asthma, that hospitalizations for

asthma among children is really primarily caused by environmental factors. And good care helps but it's primarily caused by environmental factors. I think the data are pretty clear about that.

And so I just think that in approving it, if we're going to be voting on approving it, the language just really should change. I would recommend that we ask for language that reflects that, that the science is pretty clear on this.

For instance, we don't know what causes asthma. We do know what causes asthma triggers and hospitalizations. And those are really things like, you know, exposure to mold and mildew in housing, dust, you know, roaches, pesticide exposure.

And really we've seen very effective interventions at the community level not related to care that have sharply dropped the severity of asthma cases that results in hospitalizations. So, just I do think the

language does matter a lot here.

DR. SAMPSEL: Okay. So I have Mike and then Ron and then Patricia.

MR. STOTO: Pretty much along the same lines. I'm looking at the rationale 1(b)(1) here. And I think that if you read that having heard our conversation yesterday and today you could see the right answer in there. But you could also read that and not understand that at all.

So I think that really is the place where a more careful explanation of the kind of things that John was just saying would be helpful. So it's an editorial comment rather than a comment about the measure itself.

MR. BIALEK: Yes, as a community measure I think it's a very, very strong measure where at the community level some action can be taken to make a difference.

The question I have pertains back to the stratification issue I mentioned

yesterday. Can -- well, with this particular measure where we're not making progress, where disparities could actually be widening and could be hidden in the data if one does not stratify, can we actually have a measure that specifies the measure is meant to be stratified in certain ways.

And the reason again I ask that is if we are looking for improvement to occur at the community level I don't think it really will occur unless we do have the stratification if that's part of the measure. If we just adopt the measure the way it is I think we could see the disparities grow.

MS. MCKANE: I guess I look at it just a little bit differently. And I understand the causal pathway is not all that well know. We know the triggers and we know the environmental factors.

But is this a measure of -- is it also a measure of how well controlled asthma

is to prevent hospitalizations?

I know, I've worked more with asthma epis trying to understand the rationale behind the emergency department metric for unmanaged asthma or poorly controlled asthma. And is this an extension or another way of trying to get at that is how well controlled the asthma is? Which is a reflection of -- as more broad factors than just care, but access to care and other factors.

MS. OWENS: So I don't quite understand the question other than I hear where you're going a little bit in terms of from an ED perspective versus an inpatient perspective.

This is an inpatient measure in terms of we are capturing those cases that resulted in an inpatient hospitalization. Meaning they're severe enough to require an overnight stay.

There are lots of places that that

overnight stay could be prevented from both a care perspective but from a patient compliance with their preventive medication, with your environment, with actual social determinants of health, with health education. There's many places that that hospitalization could have been prevented.

Now, on the ED portion of it you could debate and I think Dr. McInerney brought this up which is maybe the ED is the appropriate place for treatment.

And that's not at all what this measure is getting at. This measure is these asthmatics are severe enough, or this case of asthma is severe enough that it required a hospitalization that could have been prevented had a number of other factors been in place.

So I'm not sure I understand your question.

MS. MCKANE: Well, I was just trying to get at that there were other, you know,

you're talking about environmental and community factors, and that maybe this measure should be changed. And I was trying to counter that a little bit with the fact that maybe it shouldn't simply because it is getting at other factors or other -- you know, it may be a reflection of care.

I guess, I mean you're saying these are admissions that could have been prevented. Are you assessing that through coding as to what could have been prevented and what couldn't have been prevented?

MS. OWENS: In terms of whether or not they could or couldn't be prevented and coming up with a proportion that could be prevented I'd have to go to the literature for that.

I do know that the literature strongly suggests that there are lots of ways of preventing asthma admissions. So, I think in the measure testing form hopefully one of those tables talked about

preventability. If not, I have a different analytic template that I could tell you preventability which gets at some of that and that's through modeling.

But again, you know, you have to look at the individual case. So I'm not -- that's a modeling exercise to come up with preventability. Other than to say the literature all points to asthma admissions are preventable.

DR. SAMPSEL: So we're going to go Arjun, Emilio and then Ron.

MR. VENKATESH: So, I think this is a good measure and I think we're almost beating up on it too much. Because asthma as a space over the last 10 to 20 years has had so much investigation and research in comparison to a lot of these other areas that we would think about these PQI measures.

To me the things that stick out are there's almost no other PQI measure where

there's a link between community changes that can be made and the outcome measure as well as health system process of care measures that can be made and the outcome measure.

And so here's a place where both at the community level and processes of clinical care can both impact the outcome.

The fact that we're going to capture social determinants of health in the outcome to me doesn't bother me that much. And the reason is that of the 600-odd NQF measures that are endorsed you have your like, say, 60 or so that are outcome measures. They are all impacted by social determinants of health.

And I think this is the choir in this room that recognizes that. And we'll be clear in the report when we say that, that the measure is not specified to be at anything but a community level. And so to include social determinants of health as

well as those health system and clinical processes in that is totally okay. And each community can use it in the way that they need to use it.

And here's a place where, like I said, the linkage between the process and outcome is really good at two levels. And I think that it's actually easy to get past kind of question 1 here, more so than other measures.

For what Pam was just alluding to regarding the denominator being hospitalizations and not ED visits, I think that relates to a lot of these measures. Not so much here.

I think that there is enough literature to suggest that when you have inpatient hospitalizations for asthma that better outpatient care can reduce full hospitalizations as well as ED visits -- that's a separate issue -- and that community changes can reduce inpatient

hospitalization.

So I'm actually okay with this denominator here unlike some of the other measures where I think a different denominator would be better.

MR. CARILLO: Yes, I want to echo some of what's been said and just add perspective.

There's a lot of evidence that's accumulated in the last 20 years that the neighborhood effect, that there's a compounding of social determinants by neighborhood, not by county, not by large MSA which drive a lot of health outcomes.

I think that we all are saying that there is significant confounding in terms of the social determinants for this and other related measures. So I think that has to be stated.

I'm not talking about a black box warning, but it has to be stated in a way that the public, the press and others who

may refer to this measure for whatever endpoint they may have that that be put into a highlight.

The other thing, I think there's another level of confounding which is the MSA. I mean, just what I best know, New York City, Manhattan is a county. New York County. And you have East Harlem which is close to particularly environmental exposure just like the South Bronx just opposite the Harlem River where you've had historically very high rates of asthma.

And then you have other neighborhoods further down south like the Upper East Side where the levels are quite low. So, I think that if you provide a measure at a county level you're totally going to miss the fine point.

And true, it's a good measure, it says a lot, but it has to be qualified in the language that we put forth.

MR. BIALEK: Arjun, I don't think I

hear anybody disagreeing with the measure. I think it's how it's specified. And then also the point I tried to make was about the stratification piece similar to what you were just saying is that without that being specified in the measure I don't think the measure -- it could actually have an adverse impact and that's what concerns me.

MS. SELLERS: So, I guess my comment is more of a technical question which has to do with from the measure information form that we have, you know, what language from that goes directly into the NQF report versus what might need to be changed.

So, if there's language in here under the developer rationale that very clearly talks about this being a function of the healthcare as opposed to the more environmental factors is there a process to change that? Or does that get changed just because we're having this conversation?

MS. SELLERS: I'll answer that. So

we don't change the specifications. We include the specifications as part of the report.

But because this is an important issue we'll make sure that -- I think, Pam, in the submission that you sent to us this is causing quite a bit of confusion. So this might be something that we may ask the developer to change to make it a lot more clear.

Because even if in our narrative of the discussion that you're having we specify it as you would like there would be a discrepancy between what's in the submission and what we put in the report.

And so I think the committee can make a recommendation to the developer to say, to make it clear to change that language in there.

MS. OWENS: And I'm perfectly, I think it's great. I actually am an asthmatic and have been for 44 years so I

understand all of these factors and where you're coming from and I don't think it's a function of the healthcare system entirely. That being said, it's all in perspective.

So, I from my hat, I see what people download on the AHRQ website. I know who I talk to. I know how we talk about this measure. And we talk about it in this broader context.

I actually -- this form to me was just a conduit to get information to you and I didn't think it went anywhere other than you all.

Now, that being said, I totally agree with you it does not reflect my thinking, does not reflect AHRQ's thinking and from a record standpoint absolutely needs to be corrected.

So, again, I don't want you to feel like that because this form says that it either reflects AHRQ's thinking or that I took this to be anything more than quickly

trying to share woith you eight measures in a period of 48 hours to get the forms together.

MR. VALDEZ: All right, I'll take some of the blame for the interpretation.

Those of us who started using this measure in the early and mid-eighties in New York City and Los Angeles County were in fact interested in looking at the primary care delivery system and had to come up with some way of beginning to look at this.

And our initial interpretations were trying to find ways in fact that affected, or that were being affected by the lack of primary care in those communities.

And this is just reflecting historical development of something that has grown to a much bigger understanding, that in fact a number of these conditions in fact have both medical components and a community component. And you know, some of the historical initial development stuff has to

be corrected with our better understanding of how each of these individual measures operate.

DR. SAMPSEL: Okay. So, I think if there's nothing else on that first section we will go ahead and ask Kaitlynn to start the vote. And we'll start with the evidence and section 1.

Does everybody have their fun little voting things? Excellent.

MS. ROBINSON-ECTOR: For evidence the vote is open. We're waiting on one more vote.

MS. MUNTHALI: Michael Baer will not be here today.

MS. ROBINSON-ECTOR: Okay, so then we have all of the votes in. Okay, for evidence we had 20 vote yes and zero vote no.

For performance gap the vote is open. All of the votes are in. Voting is now closed. For performance gap 17 voted

high, 3 voted moderate, zero voted low and zero voted insufficient.

For high priority the vote is now open. All of the votes are in and voting is now closed. For high priority 20 voted high, zero voted moderate, zero voted low and zero voted insufficient.

DR. SAMPSEL: Okay, Jacki, do you want to introduce this section and we'll open for discussion?

MS. MOLINE: Sure. So in terms of reliability this is -- the numerator is discharges for patients 2 through 17 with an ICD-9 code for asthma. The time window is a year. This excluded children with other conditiosn that might have pulmonary complications.

And the denominator was discharges -- the denominator is -- it's discharges from the hospital. The denominator is only within the asthmatic population. The only issue is that they had to have a preexisting

diagnosis of asthma and be coded as such at the time of admission for them to be in the denominator. But it's all based on hospital-based data.

MS. ASOMUGHA: So, do the diagnoses or the patients, this is including folks who have actually died after admission for asthma?

MS. MOLINE: They would have had a discharge diagnosis of asthma death. Oh, yes.

DR. SAMPSEL: Margaret?

MS. LUCK: So I think this is similar to the measure we talked about yesterday afternoon where the numerator is the discharges of patients from a certain metropolitan area. And the denominator is the total population in that age group in that metropolitan area.

DR. SAMPSEL: Other comments, concerns, questions about reliability and overall scientific acceptability before we

vote? Okay, Kaitlynn? Oh, I'm sorry, John?

MR. AUERBACH: I guess I just wonder if AHRQ has any explanation for the regional disparities. What's going on in the West that seems to be so much better?

MS. OWENS: Well, there is something to be said for the environment in the West. I mean, people have been known to move to Arizona and New Mexico who are asthmatics.

The urbanicity. Actually some of the issues that come up in New York City are not as prevalent although we do have obviously large urban areas in the West. But how many large urban areas in the West do we have relative to the Northeast. So I think there's some environmental factors coming into play.

And I think basically you would need to look at the regional aspects combined with the environmental aspects. And in fact, CDC is looking at this issue with respect to weather changes, you know, the

environment at large. So I think there's quite a bit that might really actually account for the regional variation.

MR. ROMANO: And if I might add something.

MS. OWENS: Wonderful, Patrick. Glad you joined us. Good morning.

MR. ROMANO: Good morning. Yes, I was also going to say that there's a general pattern of practice differences.

So, even if you look at overall measures of hospital discharges per capita they're lower in the western United States than in the eastern United States.

And this is perhaps partially due to differences in training, differences in practice.

It may also reflect the higher penetration of large managed care organizations such as Group Health and Kaiser Permanente in the west coast states.

And of course some of those managed

care organizations that are closely aligned with medical groups have implemented aggressive programs to keep patients out of the hospital through better primary care management of chronic diseases.

So there's a variety of potential explanations. But this is an ongoing phenomenon that's been observed for a couple of decades.

DR. SAMPSEL: Okay. I don't see any other questions or comments, so Kaitlynn?

MS. ROBINSON-ECTOR: This is for reliability and the vote is open. All of the votes are in and voting is now closed.

For reliability 18 voted high, 2 voted moderate, zero voted low and zero voted insufficient.

DR. SAMPSEL: And any comments, concerns about validity? Jacki, did you have some comments?

MS. MOLINE: No, I think the measures are valid. They're using

standardly accepted discharge data.

DR. SAMPSEL: And discussion on validity of the measure.

MR. FRANCE: Since we have our experts here I'd like to hear more about the construct validity testing that you've done for these ACSC measures in general where you have these models that try and demonstrate that the rates of these events are aligned with things that are thought to demonstrate the health system's infrastructure.

MS. OWENS: Patrick, do you want to talk about why these models were selected for validity? Patrick, are you on mute?

MR. ROMANO: Yes, I was, thank you. Yes, if you could give me a second to pull up the results table.

MS. BURSTIN: Patrick's in California so it's 6:30 in the morning for him.

MS. OWENS: So Patrick, this was the table, and correct me if I'm wrong, that

you're speaking of in terms of the volume, the reservation quality, the model that was provided.

MR. FRANCE: I think it was a model that looked at density of physicians, insurance coverage, beds available and it was posed as a demonstration of the construct validity of the metrics as being aligned with health infrastructure, healthcare delivery infrastructure systems.

And since it plays out with all these ACSC measures we'll be talking about today I thought it would be helpful to hear a bit more about that model since it's the key piece that seems to argue that these measures are indeed reflective of the healthcare infrastructure.

MR. SPANGLER: It's table 3 in the testing form. I think it's page 7 of 12.

MS. OWENS: The reason I'm letting Patrick answer this is because as part -- under the previous contract they're the ones

that made some decisions about why these models were developed in particular. I can speak to them after Patrick if need be. I have a sense of why they were chosen.

MR. ROMANO: Right, okay. Thank you for cluing me into the location of what you're looking at.

So, the notion is that these measures obtain their construct validity in part from their relationship to primary care resources that are available in local communities.

So, we would anticipate that if an area has greater availability of physicians, particularly I should add a greater availability of primary care physicians which unfortunately is not tested here, then the hospitalization rates for the condition should be lower.

On the other hand, if there's excess capacity in terms of having a relative excess of beds those beds may be utilized,

there may be more of an incentive to fill those beds by increasing hospitalization rate.

The poverty status and insurance status are of course markers of SES as we've discussed.

I think population density is difficult to interpret. In some cases it's a marker of urbanity and therefore it may be a marker of geographic proximity to services. In other areas it may be a marker of poverty or over-population and therefore more limited access to services or constraint in supply. So, after adjusting for other factors.

So, in this case and in most cases there is a significant association between the F1 variable which represents physician supply and the outcome variable.

The prior performance is a marker of the same area's rate in the previous time period. So not surprisingly there is

substantial consistency over time in area level rate and that's reflected in the lower half of table 4 with the high-parameter estimate of 0.72.

So basically what that is telling you is the single strongest predictor of the current rate in an area is the prior period's rate in the same area.

After adjusting for prior performance the impact of physician population ratio diminishes. That's reflected in the two estimates of F1 dropping from a statistically significant estimate in the first half of the table to a non-significant estimate in the second half of the table.

Does that explain the situation?

MR. FRANCE: It does. I think it was more the high-level conversation about these measures themselves.

It looks like you create an overall F1 construct variable and then an F2

construct variable between high-quality outpatient care and then market competition. And then put those in the model and look at their impact.

So, given in general that some of the construct validity of the measures are based on these models I just wanted to hear a little bit more about them. And why you chose these variables, and how confident you are that they're indeed reflecting of the care delivery programs that are in place.

MS. OWENS: And so these --

MR. ROMANO: Yes --

MS. OWENS: Go ahead, Patrick.

MR. ROMANO: I was going to say yes. I mean, thank you for the opportunity to explain it. I have to admit it is a little bit obscure.

And because of the way -- what unfortunately I don't think we provided here was a clear explanation of the F1 and F2 construct. So, that's missing, so I

apologize for that.

MS. OWENS: So, F1 and F2 constructs are derived from the area resource file. And the area resource file is at the county level. And that's how we could do this.

I will have to go back to -- Battelle is the contractor that actually ran these models and created these constructs. Patrick is actually at UC-Davis. He's not with Battelle.

So, for additional explanation in terms of the forms I think it would be worthy, exactly what Patrick is talking about, is how were these constructs derived specifically. That would just give you -- you understand the concept. But if you want to know analytically how they were derived I can get that information.

Battelle is no longer the AHRQ QI contractor.

DR. SAMPSEL: Jason.

MR. SPANGLER: So, I have a question

specifically about this measure and this construct. The coefficient was much less in this one compared to the other measures we're going to be looking at including the one yesterday.

And is that reflective of -- that these variables within this construct have less of an impact on this measure? Going along with what we've talked about, the environmental factors and other factors, other variables that aren't in this construct. Is that the explanation for that? Or is there any other explanation?

Because this coefficient is in the 0.7 and change range where the other ones are like 0.9, low 0.9, 0.8. So I'm just wondering if that's the explanation, or are there other explanations for that.

MS. OWENS: You're speaking of the prior performance?

MR. SPANGLER: Yes.

MS. OWENS: I would explain it as

prior performance plays a huge role in all of the measures. Yes, for the asthma measure. Less so if it's relative to some of these other factors.

DR. SAMPSEL: Okay. I believe we're ready for a vote on validity.

MS. ROBINSON-ECTOR: This is validity and voting is now open. All the votes are in and voting is now closed. For validity 10 voted high, 10 voted moderate, zero voted low and zero voted insufficient.

MS. MOLINE: So the feasibility of this. It's very feasible. It's readily available data sources. And it is using hospital discharge data. So the group felt that it was a very feasible measure.

DR. SAMPSEL: Any discussion on feasibility? Okay, Kaitlynn.

MS. ROBINSON-ECTOR: For feasibility the voting is now open. All the votes are in and voting is now closed. For feasibility 19 voted high, 1 voted moderate,

zero voted low and zero voted insufficient.

MS. MOLINE: So in terms of usability and use I think this is a measure that is being used as was pointed out and has been used for 20 years now. I think that there are some areas where zip code level data are available and it is being used.

But on a macro level, looking at it at a county and in urban, rural parts of the country versus other parts of the country it is a usable measure and is -- there were no issues related to that. And it was something that could be easily tracked over time.

It was noted that there are many factors out of the control of the hospital. And the environmental issues. And the regional disparities.

There was some concern about the flatness of the measure. But overall it was felt to be a usable, highly usable measure.

DR. SAMPSEL: Any discussion on usability?

MR. BIALEK: Yes, I'd like to go back to the concern I raised before. If one looks at 4(c) benefits outweigh evidence of unintended negative consequences. I think there can be negative consequences.

And yes, the measure is being used, but we're not making progress. So I question if the measure is specified correctly for what it is we're trying to achieve.

And again, I would go down to the need to further -- the measure should specify it needs to be stratified in certain ways.

DR. SAMPSEL: All right, Kaitlynn.

MS. ROBINSON-ECTOR: For usability and use and the voting is now open. We're just waiting on one vote. All the votes are in and voting is now closed.

For usability, 12 voted high, 7

voted moderate, 1 voted low and zero voted insufficient information.

MS. LUCK: I would just like to ask if the developer could update some of the references in the measure information form. I noted that in one section the most recent reference is 2009. In another it's 2008. When I think how much work has been done in this area.

DR. SAMPSEL: Okay, so before we go to the overall suitability for endorsement I think, Pam, you've probably taken a few notes.

So, just to make sure that the committee has been heard and they will also be reflected in the overall notes on the measure. But any other final comments, considerations, discussion items before suitability for endorsement?

Okay, Kaitlynn.

MS. ROBINSON-ECTOR: This is for all suitability for endorsement. Voting is now

open. We're just waiting on one more vote.

All the votes are in and voting is now closed. For overall suitability for endorsement 19 voted yes and 1 voted no. So, for measure 0728 Asthma Admission Rate PDI 14 passes.

DR. SAMPSEL: Renee?

MS. FRAZIER: I just want to make sure from a process standpoint. There were a lot of great comments made. And I voted in the notion that much of that will be included in the conversation. I just want to make very clear. Okay.

DR. SAMPSEL: Okay, so we're now going to move to the front of the workgroup summary PDF and start with -- actually, still the PQI measures and this time start working through some of the diabetes measures.

And the first one is 0272 Diabetes Short-term Complications. And we'll start with comments from Pam or Patrick.

MS. OWENS: So, diabetes is the leading cause of hospitalization and at considerable cost. Over $174 billion annually are spent on diabetic hospitalization. And the diabetes epidemic continues to rise.

There is evidence that complications can be prevented with appropriate management in primary care.

The next four indicators you'll be reviewing are all related to diabetes. Two are short-term complications, two could be considered long-term complications.

I think it was brought up in the workgroup and I just want to highlight that AHRQ is actually considering some of the workgroup's suggestions, although of course it would not have been for this submission because it came up in the workgroup.

Anyway, so PQI 1 which is the first one, the short-term complications. And then PQI 14, making sure I get the name right,

the uncontrolled diabetes admission rate could be combined. And we do recommend -- in the assessment you'll see that we recommend that they're reported together or paired so that you get the complete picture in terms of admission rate.

And this was again brought up in the workgroup is if you combined PQI 3 and PQI 16 which are the long-term complications that those could be combined in a composite.

So, while that is not actually what's on the table with respect to endorsement I just wanted to let you know that I heard what you said in the workgroups. We are actively thinking about that. But what you are considering today is each of these indicators individually because that's as it was endorsed initially.

Are there any other questions? Patrick, did you want to add?

MR. ROMANO: Well, I can just add that I found the technical report related to

the construction of the construct variables if the committee would like more information.

MS. OWENS: Thank you.

MR. ROMANO: Right. So, those construct variables were constructed by factor analysis. And it's a principal factors unrotated method.

The F1 factor is essentially dominated by physician density per capita. Again, I think other researchers have used primary care physician density per capita and had actually found stronger results. But there's a negative loading of physician density and there's a positive loading of hospitals that supply per capita which fits with the hypothesized framework that having excess hospital beds in an area will tend to lead to more admissions. Having fewer physicians per capita in an area will tend to lead to fewer admissions.

In addition, there is a loading on

poverty which is consistent with the socioeconomic issues that we've previously discussed. The loading of population density is mildly negative, but smaller than the loading of physician density, 0.45 versus 0.25 negative, indicating that people living in more densely populated urban areas have better geographic access to primary care resources, urgent care centers, other facilities that may help to keep them out of the hospital.

So that's the construction of the factors. So factor 1 as I mentioned is dominated by physician density with the negative loading hospital density, with a positive loading.

Factor 2 is a weaker factor that basically has the inverse relationship but it's dominated more by socioeconomic factors.

DR. SAMPSEL: Okay, thank you, Patrick. I will now ask I believe John and

Jane were the reviewers. And John, did you want to lead the discussion for this measure? Okay.

MR. AUERBACH: Jane and I will do it together. But I'm happy to do that. I'm sitting next to an endocrinologist. I think there are many tabs I will defer to her.

But yes, you know, the discussion so far I think has highlighted the main points. The data source here is the administrative hospital discharge claims. It's been -- this is a review of a measure that was originally approved in 2007. It's been in widespread use. And the level of analysis is at the county, city, state, or national levels.

DR. SAMPSEL: Were there any comments, discussion on the workgroup about evidence?

MR. AUERBACH: Well, I would say here, I do think there what we know about diabetes, type 2 diabetes, not type 1

diabetes, is that it is a disease that the prevalence of which has been on the increase since the nineteen eighties in a steady and consistent manner.

And that it's correlated -- we know what it's correlated with. It's correlated with eating habits and lack of exercise. And so it is an illness that is -- type 2 is largely at the level of prevalence now because of behavior factors and environmental conditions.

So, I think like our earlier discussion causation here is due to -- causation of type 2 diabetes is not due to poor healthcare delivery, it's due to environmental factors and conditions and that's worth noting. So treatment of existing diabetes for reduction of symptoms avoids some hospitalization, definitely benefits from access to high-quality care. But around the issue of causation and the fact that we've got rising hospitalizations,

that -- the background noise there is related to the obesity epidemic. So it's just worth noting because there are from the perspective of approaches that can be taken to reduce hospitalization it's another one of these set factors where social determinants are an important consideration.

MS. CHIANG: John, can I -- so first of all, I want to apologize because I actually missed the working group calls. I actually -- I'm fairly new so I attended the dental calls because I thought that was my working group.

So there's a couple of things that -- so I'm kind of starting from the beginning. And I had a couple of questions for you about the measures. Because the short-term complications, while diagnostic ketoacidosis is part of it the main reason why people are hospitalized are due to hypoglycemia and hypoglycemic seizures. So I was wondering why that wasn't captured.

And it's particularly true in older people.

And John, I hate to do this but one of the things that we really want to dispel is the notion that type 2 diabetes is caused by behavioral issues.

We don't know exactly why there's this increased rise. We assume it is parallel to the obesity epidemic. But we don't think that's the only cause. There's social determinants. There's epigenetics. There's other reasons for this that we haven't really understood. So I do want to say that yes, that's part of it, but that's not the sole factor.

And we really are trying to de-emphasize that because then otherwise if you're hospitalized for all the short-term complications of diabetes it is a less reimbursable item because it's tied to behavior. And we don't want that to be the message. Is that a fair assessment?

So that's something that I also want

to capture in part of this discussion.

MS. OWENS: Can I have Patrick respond to that in terms of -- Patrick?

MR. ROMANO: Yes, I'm just double-checking the specification. I'm not sure that I'm understanding your question about the hypoglycemia.

DR. SCHREIBER: So I think since he missed the call I think I'll recap what I heard Patrick say on the call which was these are a family of measures.

And so you have all -- in the end you have all the hospitalizations for diabetes kind of parsed out in these next four measures. And so hypoglycemia I think would fall into the one that I have which is number PQI 14 which would be -- because it wouldn't fall under the other ones. So, it would fall under uncontrolled diabetes admission, none of the above.

And I heard this in the call that there is some misclassification in these

measures because of that. So, it's based on what people are admitted for.

But the general thrust and rationale that I heard and which I believe is true is just that -- is what John I think said but very kind of glossed over it is, you know, diabetes if it's recognized and managed can be managed as an outpatient pretty much by and large. And the hospitalizations can be prevented.

And so these are outcome measures with a rationale. And that's the rationale that I heard. And I think it's true mostly. You know, not 100 percent true, but true enough for a measure or a set of measures.

So this one is on ketoacidosis, hyperosmolarity or coma. Certainly those are preventable almost completely if it's recognized and not sprung on inadvertently with some other acute event that occurs in the patient.

So, I think that's the rationale for

this measure. And then the other measures have sort of specific rationales that they have.

MS. CHIANG: So thank you, Marcel. I agree with you and I agree that this is preventable.

MR. ROMANO: And just to clarify if I might. So the 250.0 codes here are referring to diabetic ketoacidosis. I'm sorry, the 250.1 codes refers to diabetic ketoacidosis.

The 250.2 codes refer to hyperosmolar date and the 250.3 codes refer to other diabetic coma. And most of that is hypoglycemic coma or insulin coma. So these are the most severe short-term complications.

And I think it is true that among type 2 diabetics most disease are related to insulin excess rather than to ketoacidosis. But there is a mixture of hyperglycemic and hypoglycemic complications within this

indicator.

What's shared is that these are the most serious of the complications where there's actually -- in the case of hypoglycemia there's hypoglycemic coma or hypoglycemic shock. Those are the terms that are typically used.

MS. CHIANG: So Patrick, are you distinguishing between type 1 and type 2? Or is this for all people with diabetes?

MR. ROMANO: No. These indicators do not distinguish between type 1 and type 2.

DR. SAMPSEL: John?

MR. AUERBACH: So again, one thing I would highlight here that I think is particularly noteworthy in terms of looking at the data are that we're seeing a very dramatic increase in terms of hospitalization for this measure. We're seeing an increase of something like in 5 years 110,000 to 150,000 hospitalizations.

So really rapid year to year dramatic increases.

And I guess I just would -- maybe I would ask the AHRQ what's your sense of that? I mean that's -- if you correlate it with prevalence, okay. You know, it's related, you can sort of map it. Maybe it's oversimplifying to say it's related to diabetes but it's pretty closely mapped with diabetes and you can see the slope slip just the same going up. Hospitalizations go up, obesity goes up. But that's dramatic.

And it would suggest I think that something is odd about outpatient management when you've got that dramatic an increase in hospitalizations, 110,000 to 150,000 in a 5-year period. Any thoughts about that?

MS. OWENS: In terms of why it's happening I don't know and I don't know that AHRQ has a stance in terms of the why.

I can tell you that DHHS across the Department we're very concerned about

diabetes. It is a high-priority condition. Clearly something is going on and we need to get a handle on it. So there's been multiple initiatives at CMS.

AHRQ has done quite a bit of work in terms of on the research aspects. Of course NIH is always doing work in this area. But looking at demonstrations about improving the care for diabetics as well as decreasing hospitalizations in terms of the why I don't know. In terms of yes, it's important, I agree.

MR. ROMANO: The only thing I would add is that CDC I think has just reported that there may be a downwards deflection in some of these hospitalization rates just in the last year with similar data from CDC's databases, National Hospital Discharge Survey which is very similar to HCUP. So we may as we look in 2012 and 2013 data might start seeing a change there.

DR. SAMPSEL: Jason, do you have

comments on evidence?

MR. SPANGLER: The section on performance gap. Shall I wait? But we're talking about it now.

DR. SAMPSEL: Yes, go ahead, since we're talking about all of these things, that's fine.

MR. SPANGLER: Yes, so sorry. Just because we're talking about. I'm just wondering, and thanks for that comment, Patrick, but I'm wondering why -- if you have any reasons why this one, the rates are getting worse whereas if you look at 274 which is PQI 3 it's kind of up and down. The one that Marcel and I looked at, 638 or PQI 14, it's actually improving. But this one it's definitely getting worse.

And one of the reasons is to see if you have any answers for that. And if you don't, that's fine.

But I'm wondering is there something that needs to be done with this measure.

Because this measure has been since 2007. And it's not helping at all. It doesn't seem like -- we're measuring this, which is great, but it's not benefitting anything. It keeps getting worse and worse and worse.

And is there some modifications we possibly need to make to this measure? Or not? Or is there just factors outside of that? Do we just keep measuring it?

I just don't want to keep measuring it and just keep getting worse and worse and worse numbers. Because what's the benefit of even having the measure.

DR. SAMPSEL: So Pam and I might -- I guess what I wanted to ask is kind of almost similar to what Jason has -- if you can comment with the PQIs and over time they come here for endorsement. But what is your typical almost through evaluation cycle and how are you monitoring the numbers for considerations of measure revision?

MS. OWENS: So annually we do look

at these measures. And we go through the codes. Are there new codes or whatever.

What you're really talking about is not so much the measure itself as much as the use of the measure and is it being used in such a way that there's uptake, that there's an effect right.

So -- there's a component in the sense if people don't use it because they don't think it's an effective measure, absolutely. But you know, we have various implementation strategies in place to try to see how people are using the measure and what the uptake is and what the net result is. I don't know where I was going with this conversation.

(Laughter)

MS. OWENS: All that to be said, I mean, these are based on ICD-9 codes, right? And so if, you know, you parse it out maybe one of the things we certainly can look at analytically, is it one set -- there's

basically three groups of codes there. Is it one set of codes that's on the rise versus others. We'd have to look at clinical practice. What the coding -- what's going on in terms of with the coding.

The measure itself is just capturing short-term complications of diabetes. This is what we have to work with from an ICD-9 code book. So, at some level you're somewhat limited.

I don't know, Patrick, if you want to say more on this issue. I'm not sure I'm quite answering your question.

MR. ROMANO: Well, I think it's a good question. And all we can really do is to speculate. I think that perhaps there are some things that AHRQ and others could do to empirically explore this more. But I might have some concern that this may reflect some up-coding. Because as I mentioned this PQI captures the most severe complications of diabetes which will affect

in some cases the MS-DRG assignment.

So, it may be that -- and this is part of why we really encourage people to look at the set of PQIs as a set. Because the extent that one is -- for example, this one increasing whereas uncontrolled diabetes is decreasing.

That may to some extent reflect a push towards more specific coding of diabetic complications, in other words, avoiding the use of the non-specific 0.9 codes in favor of more specific codes that often drive a higher cc level and thus a higher payment.

So, I would be cautious about over-interpreting. We could potentially do some empirical analyses to explore whether if this is true, if we're seeing up-coding over time then we might expect the other markers of the marginal severity of these complications to be decreasing over time.

In other words, patients should be

staying less in the hospital. There should be lower hospital charges over time. But I would have some concern that what you're seeing as you're comparing these different indicators may reflect changes in how they're coding.

Of course, the hospital can't get away from coding some type of diabetes complication when they admit something for a diabetes-related problem. But they may pick and choose which particular codes they put first.

Not sure if that helps but it is an informed speculation.

MR. SPANGLER: That's helpful, thank you.

DR. SAMPSEL: Jane and then Eric.

MS. CHIANG: So I think part of it is actually a true epidemiologic phenomenon. Where we are seeing increased incidence of -- - well, both incidence and prevalence of both types of diabetes. And we are seeing -

- with the greater numbers we are seeing greater admissions.

The other factor is a lot of primary care physicians don't know how to recognize the issue. And this has been part of the problem as well. Just from, you know, from my perspective.

So I don't know if you can just say it's necessarily the tool, but I also think that this reflects what we've seen in practice as well.

The other thing I was going to say is just like with all the others the economic analysis, if you're going to use that as an argument there is a paper that we published last year which reflects the 2012 economic cost analysis of diabetes.

And that reflects a $245 billion net cost. And then the hospitalization cost is $176 billion. So you may want to use those numbers.

DR. SAMPSEL: Eric, did you have

something?

MR. FRANCE: Probably just a question about further development of metrics knowing that the prevalence is going up with diabetes across the country and that's embedded in this per 100,000 population measure, whether you have other metrics, where you're looking with the denominator being people with diabetes.

So that one can look at among people with diabetes is the health systems infrastructure doing a good job of managing them.

MS. OWENS: So, excellent question. And as part of the QI program one of the things that we have looked at is can we get county-specific rates of diabetes. So, you know, what's the prevalence of diabetes.

And compare that to using a variety of things. And we actually are using CDC data to try to look at that a little bit more in-depth.

The other thing -- so to answer that specific question, ongoing. Okay? No definitive answer yet.

The other thing is that these PQIs is they are being adopted by various programs whether they be CMS programs or state programs. Some of them are trying to see whether or not they use this numerator in one of these four, but change the denominator such that the denominator reflects the beneficiary population with diabetes. And then they do a lot of reliability and validity testing to see if in fact that indicator is still stolid. So, it's an adaptation of the PQI. That's not what's before you for endorsement, but yes, that's being done in terms of being looked at and the testing is underway. Does that --

MR. FRANCE: I think that sounds great. You're always sort of stuck. It actually sort of raises more questions when

you see the results. Is this because of the prevalence is increasing and that's why it's flat, or is it because our infrastructure isn't very strong. And so you sort of go into those next level of questions to try and get an answer to that. So it's the limitation of the value of the current metric when it's looked at alone.

DR. SAMPSEL: Jane, did you have a final comment before we go to vote? Okay.

So, Kaitlynn, I think we're ready to vote. We have, Arjun, you're recusing yourself, correct? And then Mike, you're recusing yourself as well due to involvement on advisory panels or technical expert panels. So, and then Tom has rejoined us so I think we're looking for 19.

MS. ROBINSON-ECTOR: This is for evidence and the voting is now open. Two votes are out. I think we've missed one person so we'll have to go back. Voting is open. It looks like one vote is still out.

DR. SAMPSEL: So that's pretty clearly yes. And so we'll move on. But basically when you click you should also see a little green light go on, so if folks want to make sure that's happening.

MS. MUNTHALI: Maybe you can do a test so that we can make sure everyone's device is working.

DR. SAMPSEL: Okay. So this next one will be a test.

We'll try again on a test. And everybody, even if you're recusing yourself should be -- you can't recuse yourself from a test. Let's go with that.

MS. ROBINSON-ECTOR: Okay, so all of them are working. There's 21.

DR. SAMPSEL: All right. So let's just continue to go through. They all worked that time but we do know that first one, unless somebody knows they said no and doesn't come to 1, whatever.

But let's go ahead and continue to

move through 1. And during our discussion we do go all the way through performance gap and priority. So Kaitlynn, if you could just take us through the two remaining votes.

MS. ROBINSON-ECTOR: For performance gap the voting is open. All the votes are in and voting is now closed.

DR. SAMPSEL: Jason, you have a question?

MR. SPANGLER: Yes, I had a question about this in the other measures around looking at performance gap.

We have these observed rates but there's no baseline. Like I don't -- to me that seems bad but is there a good rate? You know what I'm saying? I mean, the rates are variable. Some are rising, some are not.

But, I mean, if all of them were, you know, 11, 12, 13, 14, 15 they'd all be worsening. But that would obviously be a

lot better.

So is there any -- I guess maybe the people who are -- maybe Jane, the endocrinologist.

But I mean, not that there's an acceptable rate but as something good versus really bad. I don't know if you guys from AHRQ know. I mean, I don't know what you would say to that.

MS. OWENS: It's a good question. I would actually defer to her.

MR. SPANGLER: Like an observed admission rate. You know, for these patients. I mean we have rates that are around 50-60 per -- for what they have. So, per 100,000. So, I'm wondering. You know, I'm sure there's some rate that's acceptable, right?

MS. CHIANG: I mean ideally it would be zero, right? Where you would educate the -- I think this is multifactorial where if you do have patients who recognize what it

is and the providers who recognize it it would be zero. But since that's not the case as low as possible.

And I think they're both bad but I think for DKA it's probably a little bit -- it's okay to have it a little bit higher than HHS because that's very bad to use your terms.

MR. SPANGLER: Got it, thanks. I like those terms, thanks.

DR. SAMPSEL: I think, Jason, that's still your card up. It's hard for me to see that far. Go ahead and continue, Kaitlynn.

MS. ROBINSON-ECTOR: Seventeen voted for high, two voted for moderate, zero voted for low and zero voted for insufficient.

For high priority voting is open. And we're waiting on one vote. All the votes are in and voting is now closed.

For high priority 19 voted high, zero voted moderate, zero voted low and zero voted insufficient.

DR. SAMPSEL: Okay, John, back to you for reliability and validity.

MR. AUERBACH: The reliability is established. And the numerator and denominator are clear. I do have a picky question about the not counting pregnant women in the denominator which I mentioned on the call. It's a minor issue but I do think that data exists and you can actually decrease pregnant women from the denominator. The state data are available for that. But that's a minor point.

And the measure has been well tested for reliability.

DR. SAMPSEL: Jason?

MR. SPANGLER: I had another question for the AHRQ staff. There was no mention through this and several others around ICD-10. And I'm assuming that's something you guys are already planning for when those changes come.

MS. OWENS: Right. I'm not sure

what was submitted in the Excel spreadsheets.

All of these codes have been mapped to ICD-10. We did post it for public comment in December and that public comment period ended on December 25. Good timing on the government's part.

But all that to be said we did not receive any comments that suggested we needed to change those mappings. Actually, Patrick and his colleague Ginger Cox are actually at the forefront of doing that mapping.

We will have a beta version -- our plan is to have a beta version of all of these indicators mapped to ICD-10 in a SAS program by October 1, 2014. Of course we have been delayed on implementation.

MS. MUNTHALI: I think there was some confusion on the workgroup. What we received in the Ecxel spreadsheet was corrupt. And I'm not sure if we got the

correct worksheet. But, that's -- we'll include that and share it with the committee once we receive that.

DR. SAMPSEL: Other discussion points on scientific acceptability?

MR. AUERBACH: I will say --

MR. ROMANO: I will say -- this is Patrick. Oh, sorry.

MR. AUERBACH: I would defer to him.

MS. OWENS: Go ahead, Patrick.

MR. ROMANO: I was just going to say that there are some changes as some of you may know in the ICD-10 CM codes. It may lead to some upward or downward movement for some of these individual PQIs.

We wouldn't expect the group as a whole to move up or down specifically because of the ICD-10 implementation. But ICD-10 as you may know offers more specific codes for some diabetic complications.

It also offers a greater categorization of what's referred to as

secondary diabetes, different types of secondary diabetes. So this may have some impact of how cases sort across these four PQIs. That's all I'll say.

But anyway, the specifications are available for comment and there will be further testing that will happen over the next year and a half.

MS. OWENS: And just to clarify that testing will take dual coded data from the medical records where the medical records have been coded in both ICD-9 and ICD-10. We will be able to determine if you're using the I-9 algorithm was this an identified case.

Then the same case, the same record has been also I-9 coded. We'll run it through the I-10 specifications and see if the same case was identified.

MR. AUERBACH: So I would say in terms of the issue of validity the specifications align with the evidence of

quality as indicated by the testing that has been done.

However, I think that as we're talking, as we're discussing it the quality of care is insufficient to explain the dramatic increase in hospitalizations. And that there are complicating factors associated with that which I think are socioeconomic in nature.

You know, we see a dramatic difference between diabetes level in Black and Latino communities, for instance, than White communities and in poor communities.

So I just think in terms of the validity there's room for more testing and looking at factors other than quality of care as correlated with this.

I hope, I don't know if AHRQ is planning on doing that but I think if there are ways of looking at that it would be useful for further considerations.

DR. SAMPSEL: Okay, Kaitlynn.

MS. ROBINSON-ECTOR: This is for reliability and voting is open. We're just waiting on one more vote. All of the votes are in and voting is now closed.

For reliability 14 voted high, 5 voted moderate, zero voted low and zero voted insufficient.

This is for validity and voting is open. All votes are in and voting is now closed. For validity 11 voted high, 6 voted moderate, 1 voted low and 1 voted insufficient.

DR. SAMPSEL: John.

MR. AUERBACH: Here as we talk about the data sources, discharge, diagnostic claims, easily and readily available electronically as well as on paper.

DR. SAMPSEL: Any other discussion points? Okay.

MS. ROBINSON-ECTOR: This is for feasibility and voting is now open. All votes are in and voting is now closed.

For feasibility 18 voted high, 1 voted moderate, zero voted low and zero voted insufficient.

MR. AUERBACH: So with usability this is currently widely in use as a measure in at least 12 states and by CMS, and has been in use for some time.

With regard to the improvement over time and the usefulness of data I think we've already discussed that, that disturbing that -- that while we're looking at this it doesn't appear to be useful in terms of improving care but nonetheless it is widely utilized as a measure of -- as a quality measure.

DR. SAMPSEL: Other discussion points on usability or questions. Jane?

MS. CHIANG: So we know that this is usable. It's been used in the past. But is it the best? And so from my information would this be considered -- so for example, if we continue using it as is it sounds

fine. I would rate it a high.

But if we were to incorporate some of the suggestions that would be better. So how would that influence my voting? So would that be a moderate or low?

MS. OWENS: No, you need to consider --

MS. CHIANG: As is?

MS. OWENS: -- as is, as when presented.

DR. SAMPSEL: Okay, Kaitlynn.

MS. ROBINSON-ECTOR: This is --

MR. ROMANO: One question, a brief point. The question was raised earlier about what's the right rate or how do we point people towards what they should be able to achieve.

There is a concept called achievable benchmarks that we sometimes recommend which basically points the 20th or 25th percentile as the empirical distribution as what should be an achievable benchmark for most areas.

So, in this case, again, recognizing the ideal rate would be zero but very few if any communities are able to achieve a zero rate.

However, 25 percent are able to achieve a rate of 40 which is less than half of the 75th percentile and less than one-third of the 95th percentile. So, that's the general approach that we use sometimes to steer people towards what should be achievable for most areas.

MS. ROBINSON-ECTOR: This is for usability and voting is now open. All the votes are in and voting is now closed.

For usability 13 voted high, 4 voted moderate, 2 voted low and zero voted insufficient information.

DR. SAMPSEL: So our next vote will be for the overall suitability for endorsement. Any additional comments, reflections, discussion points before we vote yes or no on this one? Okay. Oh,

Robert.

MR. VALDEZ: I just wanted to say that it's important as we began our discussion to recognize that this is one of a set of measures. This one taken alone probably doesn't tell us a lot, but taken in combination with others is probably the one that gives us some answers to some of the questions that have been bandied about.

MS. ROBINSON-ECTOR: This is for overall suitability for endorsement. Voting is now open. We're waiting on one more vote. All the votes are in and voting is now closed.

For overall suitability for endorsement for measure 0272 Diabetes Short-Term Complications Admission Rate PQI 01 19 voted yes and zero voted no. The measure passes.

DR. SAMPSEL: Okay, great. So we're now going to move onto 0274 and in our workgroup summaries if folks are looking at

that that's page 7. And discussants, I think Emilio, are you ready to lead this? He is.

And I think as with the past measure Arjun and Mike, you are both continuing to recuse yourselves, correct? Yes?

MR. STOTO: I'm not recusing myself from this one.

DR. SAMPSEL: Oh, you're not. Okay. We'll put you back on the list. So go ahead, Emilio.

MR. CARILLO: Yes. As has been said this is part of a suite of measures which taken as a whole has increased value.

It's an outcome measure. The 0274 which is the diabetes long-term complications admission rate PQI 3.

It's sole source. It's administrative claims which we've discussed has issues here and there but it's something that we see throughout all of our measures.

It is population-based and there is

a very well established path between diabetes and long-term complications associated with microvascular damage.

We have the United Kingdom Prospective Diabetes Study, famous UK PDS study that has shown great relationship between the two.

And in terms of the evidence, I mean there's a number of evidence-based guidelines that are based on this pathway. So clearly there's good evidentiary value to this.

Not that it makes a significant difference given how we are choosing principal diagnosis across the board, there could be a significant underreporting of diabetes long-term complications.

The National Hospital Discharge Survey, the CDC in 09 showed that there were 688,000 discharges with a principal diagnosis of diabetes mellitus as opposed to 5 and a half million if you look at the top

seven discharge diagnoses. But understanding that we do have the standard of the principal diagnosis.

The issues regarding social impact, the MSA, level of analysis, zip code versus greater county, those issues apply just like they did in all the previous discussions.

In terms of the specifications I would again raise an issue that Ron and others have raised which is stratification being of value.

In determining predictive value the measure developers just looked at age and gender as covariates and perhaps racial and ethnic covariates and other SES covariates would give further value.

So, again, let me stop there and get any other comments in terms of evidence.

DR. SAMPSEL: Any discussion items on evidence? Jane?

MS. CHIANG: This is a question for AHRQ. So, if someone is discharged from the

hospital with let's say a heart attack but they don't co-code for diabetes, is that captured?

MS. OWENS: So, this is a principal diagnosis of diabetes. And so, that particular instance where you're telling me that the principal diagnosis after all is said and done, that the reason for the hospitalization was the heart attack, that would not be captured here. Right?

MS. CHIANG: Because that happens a lot, where a lot of times people are -- the primary condition is not captured. And I think for the sake of this it kind of influences the way I see the evidence.

MS. OWENS: Patrick, can you inform the group? Because I know you were involved in the development. Was this taken into consideration? And what the thoughts were?

MR. ROMANO: No, that's exactly right. I mean, these are not measures of the total population burden resulting from

diabetes. So, clearly if you were to tally all diabetes-related hospitalizations in the United States the total would be much greater. So these are hospitalizations that are specifically linked to identified diabetes complications, neurologic complications, diabetic gastroparesis, so forth.

And therefore it is an undercount given that the cardiovascular complications of diabetes are of course what take the largest toll in terms of deaths.

If we were to include those the numbers would be even greater. But of course we have to then figure out some way to apportion which of those, for example, diabetes-related MIs are actually linked to diabetes.

In the case of the long-term complications that are included here we rely on physician labeling and the coding to say that diabetic gastroparesis, for example, is

ipso facto a result of poorly controlled diabetes over time. Similarly for diabetic neuropathy.

So, it's just a matter of sort of separating those complications that are more intrinsically linked to the diabetes versus those that are multifactorial where diabetes is clearly a contributing factor but it's a little bit more difficult to assign the complication definitely being due to diabetes. Does that make sense?

MR. CARILLO: Perhaps before voting I could just make a comment on performance and high priority. Or we'll discuss that later?

DR. SAMPSEL: Sure, that's fine. Sure, go ahead.

MR. CARILLO: Just, again, nothing earth-shaking. In terms of performance gap there is a very solid robust gap. The developer looked at HCUP data from 07 to 11. Very, very robust distribution of

scores, 5th to 95th percentile.

And also ethnic and racial minority disparities exist and well documented. They cite 51 studies and there's probably more than that.

And I -- you know, our group in this discussion felt that this is a high priority, well demonstrated, pathophysiologic chain, and affecting large numbers. A significant cause of morbidity and mortality and major resource consumption.

DR. SAMPSEL: Great. So, hearing anything else? Tom.

DR. MCINERNEY: Now, I'm a little confused. Because I worry about this coding. As someone mentioned it has to be coded as a complication of diabetes? What happens if the code, the first code is renal failure? And the reason for the renal failure is because the patient has diabetes. Now, how is that coded?

And I think in some instances -- I don't understand the coding. Do you understand how that works?

MS. CHIANG: It really is physician-dependent. But one thing that frequently happens is diabetes is not captured as the primary etiology behind the cause.

So, we know that a lot of deaths, for example, are due to myocardial infarction related to macrovascular disease. But that is not captured. It just says myocardial infarction.

So it's a problem because you can't capture the enormity of this disorder because the main reason for the death is not identified.

DR. MCINERNEY: And then to further complicate matters what happens if the diabetic patient is also a smoker? And then how does that color it?

MR. SALIVE: Yes, so I think this is why that's not a big deal. You know, the

things you're talking about have multifactorial causes. And we can't sort it out in claims data. It's really a blunt instrument. We're looking for preventable complications here which I think if they say that's why they were admitted, it's the principal diagnosis, okay, enough.

Because there's one million heart attacks a year and sure, some fraction are due to diabetes, but we don't know what fraction and we don't know which admissions. And that's not the focus of this measure. So, you know, very interesting discussion.

MS. CHIANG: But I actually disagree because I think that if we're the health and well being group that it is a preventable measure.

So, for example, the UK PDS, the Diabetes Control and Complications Trial all showed that if you can reduce your A1c -- they did show that in the DCCT trial and the UK PDS. It did show that if you decrease

your A1c to have more intensive management you can have improved retinopathy decreases, all of those factors decrease.

So I think it is important. Whether we can do something about that, I think that's a different story. But if you can control the disease you can control long-term outcomes.

But that's a separate topic I think from what we're trying to do here which is really to say is this measure effective.

MR. STOTO: I would build on that last thing. The question is is this measure good. We're not trying to sort out whether every one of these cases is or is not due to diabetes. The question is do they assist us in making comparisons over time and between locations that are helpful.

And to the extent that the fractions of cases missed or reported incorrectly are relatively constant I think the measures work.

MR. CARILLO: Just to clarify something that maybe is clear already, that we're talking microvascular damage. So, heart attacks are not part of that. That's macrovascular.

MR. ROMANO: Just one clarification of coding rules.

So, coding rules do specify that if a patient is admitted with acute or chronic kidney failure and they have a diagnosis of diabetes then the default position for the coder is that it gets coded as a diagnostic complication, as diagnostic nephropathy.

That is not always true but for all diabetic complications, it's not true for cardiac complications, but it is true specifically for renal complications. The default position for coders is to assume that nephropathy in a diabetic is diabetic nephropathy and therefore it gets captured here.

Of course there may be some cases

where the physician fails to diagnose diabetes at all on a record of a hospitalized patient. That hopefully is not too common but it could happen.

DR. SAMPSEL: Thank you, Patrick. Go ahead, Kaitlynn.

MS. ROBINSON-ECTOR: This is for evidence. And voting is open. All the votes are in and voting is now closed.

For evidence, 19 voted yes and 1 voted no.

This is for performance gap and voting is now open. We're waiting on one more vote. For performance gap 15 voted high, 4 voted moderate, zero voted low and zero voted insufficient.

This is for high priority and voting is now open. We're waiting on one vote. All the votes are in and voting is now closed. For high priority 18 voted high, 1 voted moderate, zero voted low and zero voted insufficient.

MR. CARILLO: Reliability. The metric of reliability is signal-to-noise ratio. And developers and AHRQ have done two sets of measurements on this. And both come in very positive. So there is good evidence to support the reliability of this.

DR. SAMPSEL: Any other comments on reliability? All right, let's go to vote.

MS. ROBINSON-ECTOR: This is for reliability and voting is now open. All the votes are in and voting is now closed. For reliability 17 voted high, 2 voted moderate, zero voted low and zero voted insufficient.

MR. CARILLO: The workgroup had concerns about the validity of this measure. The very fact that it takes decades to develop the microvascular damage, it's significant.

This is very different than the measure of asthma which is something that happens acutely or subacutely. This is different than what happens in short-term

measures of diabetes complications.

We're talking about complications that take 15-20 years, maybe more to develop.

So what are we measuring? When we measure -- when we get a measure of these long-term complications what if we have a community, say, in New Mexico where a new Kaiser comes in, rolls in and over 10 years develops a whole set of primary care opportunities. And you may be measuring what was there before Kaiser came in, 10 years before, 15 years before.

And conversely, if you have an area that loses their main industry, sort of a Detroit effect and over 10 years the industry is gone, the primary care is gone, and the measures that you get may reflect life before the industry left.

So I think that there has to be real care in how we use this measure. Again, as a suite of measures that possible

confounding is diminished. But nevertheless I think that there's something oxymoronic about this.

DR. SAMPSEL: Other comments about validity? Okay, Kaitlynn.

MS. ROBINSON-ECTOR: This is for validity and voting is open. All the votes are in and voting is now closed.

For validity, 4 voted high, 10 voted moderate, 4 voted low and 1 voted insufficient.

MR. CARILLO: The feasibility was discussed by the group and all felt that there's feasibility at all levels, public health departments, researchers, ACOs, HMOs. And it's a positive statement.

DR. SAMPSEL: Any discussion around feasibility?

MS. ROBINSON-ECTOR: This is for feasibility and voting is open. We're waiting on one more vote. All the votes are in and voting is now closed.

For feasibility 19 voted high, zero voted moderate, zero voted low and zero voted insufficient.

MR. CARILLO: And finally, in terms of usability and use this has been around since the eighties. And my friend here from L.A. -- you didn't use this measure? Well, I take that back. This has been around for a long time.

(Laugher)

MR. CARILLO: I was trying to give him credit for something. And it's been very useful.

DR. SAMPSEL: Any other discussion about usability? Kaitlynn.

MS. ROBINSON-ECTOR: Voting is open. There's one more vote left. One more vote still out.

So for high, 10 voted for high, 7 voted for moderate and 2 voted for low.

DR. SAMPSEL: Okay, next vote is overall suitability for endorsement. Any

additional discussion items? Anything else folks would want to comment to AHRQ? Okay, let's vote.

MS. ROBINSON-ECTOR: Voting is open. All the votes are in and voting is now closed. For overall suitability for endorsement for measure 0274 Diabetes Long-term Complications Admissions Rate PQI 3, 18 voted yes and 1 voted no. So the measure passes.

DR. SAMPSEL: Great. Two more PQIs. We're going to change things up a little bit here and go to amputation. And I'll ask as before if Pam, if you will make some introductory comments. And then I believe our lead discussant is Patricia for this one. Great. So, Pam?

MS. OWENS: I don't have any additional comments other than what I said at the beginning which is the potential to bring this in with the other measure.

MR. FRANCE: Is there time for a

break, Sarah? I think 10:45 had us at a break.

DR. SAMPSEL: Sure, we'll go ahead and take a break. We were going to try to get through but why don't we go ahead and give folks some -- 10 minutes to stretch your legs.

(Whereupon, the foregoing matter went off the record at 11:09 a.m. and went back on the record at 11:15 a.m.)

DR. SAMPSEL: Okay, we're going to make just a real minor change to the schedule. We're going to go ahead and actually move 0638 the uncontrolled diabetes admission rate up just because it is so similar to the other measures.

And perhaps we can make some efficiencies out of discussions we've already had. But we do still need to go through the process and have those discussion items where warranted and where committee members want to bring up

considerations.

We'll then do the amputation and diabetes measure, and then go ahead to the NCQA breast cancer screening measure.

We're just trying to adapt to being a little bit behind schedule and knowing that some of the measure developers will have to leave between 12 and 1.

So with that 0638 -- I promise, Patricia, we'll come back to you -- 0638 was Marcel and Jason. And I don't know which of you was going to lead.

MR. SALIVE: I'll take it since Jason left the room.

(Laughter)

MR. SALIVE: He must have heard wind of this plan.

So this measure is for principal diagnosis of diabetes without -- as I said I think about an hour ago without mention of the other complications that we've just gone through. So, it is a complement to those

last two measures we just looked at.

And it's described on the screen. I think it's still an ambulatory-sensitive condition, avoidable hospitalization.

The rationale being that the management of diabetes is done as an outpatient and does not need to be done in the hospital. So I think that's the -- it's an outcome measure with the rationale.

The measure is pretty stable I think over time, but it does show variation with respect to certainly age. It's very increasing with age considerably and highest in the 75-plus age group.

There is a performance distribution score, distribution I guess as mentioned by the developers. So the 25th percentile is 8.6 and then the highest, it gets up to like 40. It's been stable so there is a gap.

I think, you know, we don't have to discuss the importance of diabetes. It's hugely important. So I think I covered the

first three right there albeit efficiently and quickly.

DR. SAMPSEL: Appreciate it. We don't want to stunt conversation at all. So at that are there additional comments, questions, considerations for this measure? And Pam, did you have anything you wanted to add? Jane?

MS. CHIANG: This is more informational but where is the information captured for those less than 18? Is that something that's captured, or is that part of this discussion?

MS. OWENS: We have a different set of measures called the pediatric quality indicators and that's -- the two that you looked at this morning were pediatric quality measures.

In terms of in the pediatric quality measure set, let's see, we do have a diabetes short-term complications rate. That has not been brought to NQF for

endorsement.

DR. SAMPSEL: Other comments on evidence, opportunities for improvement, gaps, et cetera? And who is recusing from this one? Arjun? No? Mike, no? Okay. Go ahead, Kaitlynn.

But Kaitlynn, we're missing Jason and Tom. Yes.

MS. ROBINSON-ECTOR: This is for evidence. And the voting is open. There's one more vote out. So for evidence 19 voted yes and zero voted no.

This is for performance gap and voting is now open. One vote is out. All votes are in and voting is now closed. For performance gap 18 voted high, 2 voted moderate, zero voted low and zero voted insufficient.

This is for high priority and voting is open. One vote is out. All votes are in and voting is now closed. For high priority 18 voted high, 1 voted moderate, 1 voted low

and 1 voted insufficient.

DR. SAMPSEL: Okay, Jason or Marcel, scientific acceptability, validity, reliability.

MR. SALIVE: Okay. So this one, again, it's a measure that's been endorsed in the past. It's coming for resubmission.

They used for reliability the signal-to-noise ratio and presented some data on that which was -- the developers have judged that as moderately reliable.

They did the construct validity testing looking at structural measures to predict and substantial evidence was presented on that construct validity.

I think on the call we did have some discussion about some possibilities of misclassification just that might affect validity as a family of measures. I think that accounts for some of the possible misclassification.

DR. SAMPSEL: Mike?

MR. STOTO: I just want to say that we have to really think about this as part of a family of measures. By itself it's probably -- it's kind of a leftover category. It doesn't have that much meaning. If that's so indicated.

DR. SAMPSEL: Eric?

MR. FRANCE: In that same line I was just thinking that with these very small rates of 18, 17 per 100,000 events in the context of these other metrics that are being followed by counties what is added by having this third metric, except for the fact that it accounts for some things that might not be otherwise be classified.

So it might be a conversation for later when we're talking about harmonization across measures. Is there a utility in tracking this over time and is it reliable in small communities given the low rates, or could it be potentially blended into the short-term outcomes, for example, as a

future metric.

DR. SAMPSEL: Arjun?

MR. VENKATESH: I think you can almost answer that question a little bit. I think the benefit of having this measure is that for measures like this where you have this risk of misclassification across categories there's always some measure reliability that you'll lose with coding drift over time, or potential gaming.

And so by basically giving yourself the last piece of the pie I think it helps the other measures that you know over time whether or not to look for things like coding drift or gaming.

MR. SALIVE: And I think the developers did present data suggesting it's reliable to communities of size greater than 15,000 population which is not too bad.

DR. SAMPSEL: Eric, did you have another comment?

MR. FRANCE: I don't know, I just

didn't make the mental connection about how coding drift or gaming is somehow helped by having this third category.

MR. VENKATESH: It would be more that if you didn't have this category then over time instead of coding things as short-term complications then they get coded in this category and completely unmeasured.

So while you think you're doing better because your short-term complication rate is declining in the absence of this measure you wouldn't see this going up. And so having this measure lets us know if people are -- instead of being in one bucket are getting coded in a different bucket.

MR. ROMANO: I might add one other point.

MS. OWENS: Go ahead, Patrick.

MR. ROMANO: Yes, it's just to point out that from the standpoint of healthcare interventions there's a second pathway that's relevant to this particular indicator

which relates to vascular care.

So there's ongoing efforts of course to reduce amputation rate by improving vascular care which would include both potential re-vascularization of patients with peripheral arterial disease as well as better treatment of lower extremity ulcers in patients who have diabetic vascular disease.

So, this is an indicator the vascular surgeons and vascular programs are particularly interested in for that reason.

MS. OWENS: So, Patrick, that actually applies -- we switched orders on you, sorry. I think you're talking about the lower extremity amputation measure.

MR. ROMANO: Oh, I am. I'm sorry. I came into the discussion at the wrong time. Sorry about that.

DR. SAMPSEL: We're just trying to keep you on your toes, make sure you're still awake.

So with that if there are no other -- oh sorry, Jane.

MS. CHIANG: So, I understand why you would have this measure, but one thing that -- because some people just admit because someone is not in control, and that's the way you would manage these patients.

I want to reiterate my concern again about capturing hypoglycemia as a specific. I heard that it was part of this but I don't see it as a specific call-out. And I think that that would be very useful.

MR. ROMANO: So I'm sorry, just to be clear, are you suggesting that there should be a separate indicator for the hypoglycemic complications versus hyperglycemic?

MS. CHIANG: Correct.

MR. ROMANO: Okay. Well, that is something that could be evaluated and tested.

DR. SAMPSEL: Any other comments? Okay, Kaitlynn, can you lead us through reliability and validity?

MS. ROBINSON-ECTOR: Voting is open. All votes are in and voting is now closed. For reliability, 9 voted high, 11 voted moderate, zero voted low and zero voted insufficient.

Voting is open. All votes are in and voting is now closed. For validity, 4 voted high, 15 voted moderate, 1 voted low and zero voted insufficient.

DR. SAMPSEL: Marcel, feasibility?

DR. SCHREIBER: Feasibility. It's a claims data measure and uses a population-based denominator from Census data. I have no concerns about that for this measure.

It's used by many -- CMS and Monarch and numerous states report it so it has high usability. And we've discussed the other measures.

DR. SAMPSEL: Other comments,

questions, feasibility? Okay.

MS. ROBINSON-ECTOR: For feasibility and voting is open. All votes are in and voting is now closed. For feasibility 19 voted high, 1 voted moderate, zero voted low and zero voted insufficient.

DR. SAMPSEL: Okay, Marcel, you just commented on usability. Were there any other comments or considerations for usability? Okay.

MS. ROBINSON-ECTOR: Voting is open. All votes are in and voting is now closed. For usability 13 voted high, 6 voted moderate, 1 voted low and zero voted insufficient information.

DR. SAMPSEL: Okay, any final comments before overall suitability for endorsement? Great.

MS. ROBINSON-ECTOR: Voting is now open. All votes are in and voting is now closed. For overall suitability for endorsement for measure 0638 Uncontrolled

Diabetes Admission Rate PQI 14, 19 voted yes and 1 voted no. So the measure passes.

DR. SAMPSEL: Okay, now we're going to go to 0285 Rate of Lower Extremity Amputation Among Patients with Diabetes. I believe Pam had already made minor comments and Patricia, you're going to kick us off.

MS. MCKANE: Okay. This measure is for admissions, any listed diagnosis of diabetes and any listed procedure of lower extremity amputation per 100,000 population ages 18 years and older.

It excludes any listed diagnosis of traumatic lower extremity amputation admissions, toe amputation admissions which are likely to be traumatic, obstructed admissions and transfers from other institutions.

It's using the same data source as this whole family has been using. As Patrick spoke out on this about the rationale and the evidence this is an

outcome measure. The workgroup agreed with the measure that the focus of inadequate diabetes management screening will prevent lower limb amputation linked to diabetes.

So we felt that this measure -- we had many of the same concerns that were addressed earlier regarding the data source, regarding the fact that there's other factors that could contribute to this that may be confounding this outcome.

It's a bit more distal than perhaps some of the others. And our workgroup summary, it does allow comparison across regions. This is increasing. We do see that measure -- to assess preventive education, outpatient care and management of diabetes and access to care where these resources are lacking since high-quality education and care management and early intervention has been shown to result in lower rates of amputation linked to diabetes.

We thought this needs to be, again, within the context or even possibly changed to a composite measure.

We also talked about using a geocoded data set to allow a more precise estimate, and particularly to pinpoint high-risk neighborhoods.

Again, just there were also disparities by income, region, gender and age that we noted. Prevalence I believe is increasing. Again, this may be a lower prevalence but it is a really extreme outcome. And I can't remember the other things we were supposed to talk about.

So I guess I will leave it to my workgroup to fill in any blanks that I may have left out, things that you don't think I covered.

DR. SAMPSEL: Jane?

MS. CHIANG: So, actually lower limb amputations, the rate has gone down tremendously over the past 10 years. So

that's the one area where there's been a significant difference. The data doesn't capture it here, but in general this is one area where I think it's like 45 percent where it's gone down.

Prevalence has gone up for diabetes as a whole but because people are aware and there's better vascular maintenance I guess it has gone down substantially.

MS. MCKANE: Okay, I was going by what was in the packet.

MR. AUERBACH: And what would be the reason that that wouldn't be captured in these rates? Do you have an explanation for that? Because this shows the rates going up.

MS. MCKANE: Right.

MS. OWENS: Exactly. So, the data that you're looking at, it would be useful for us to look at why there's this discrepancy. Because in fact it looks like it's doing the opposite. I don't know why

that is.

MR. FRANCE: I would imagine it's the rate per 100,000 population versus a rate per 100,000 diabetic patients. So you're seeing the reduction in diabetics. And then -- which is staying flat because the incidence of diabetes is high.

MS. OWENS: Right.

DR. SAMPSEL: Arjun?

MR. VENKATESH: The other thing is that all these measures are visit rates. So if, for example, somebody had four hospitalizations for each being one to amputation it's counted four times versus previous event may have been multiple toes amputated at once. And so for all the PQI measures these are events, not patients.

DR. SAMPSEL: Other comments in the evidence area? John.

MR. AUERBACH: So I guess if in fact this isn't the best measure of what's actually happening, I guess I'm just raising

a question about that. Is there something we're missing in terms of this that might allow us to better understand what these trends are?

MS. OWENS: Patrick, I'm going to defer to you because perhaps -- I'm not actually familiar with what you're saying. I'm not doubting it, but I'm just not familiar with it.

Patrick, do you know, have you discussed some of the discrepancy that may be coming out of the endocrine societies versus what HCUP is showing?

MR. ROMANO: Yes, I think that the discussion -- I would agree with the discussion thus far.

So there are two countervailing trends. One is the increasing prevalence of diabetes and the second is a decreasing amputation rate among patients with diabetes. So those are basically washing out. So we're seeing effectively no change

in the rate of this indicator.

So it ties into previous discussion of -- as we move towards potentially adjusting all of these diabetes-related PQIs for diabetes prevalence at the community level that would fix the problem.

But of course the problem is getting sufficiently reliable and valid estimates of diabetes prevalence at the individual community level. So it works well for big cities. It becomes more problematic for a lot of our larger states that have large rural populations.

The other thing I would say is that we have had some discussions and certainly would be interested in the steering committee's input on this.

So some have argued in favor of removing toe amputation from the specification here for just the reason that was mentioned, that in some cases the toe amputation is something that has relatively

little functional effect but forestalls progression and prevents a more functionally significant proximal amputation.

In addition, some patients do come back for multiple amputations of multiple toes and that may be seen as inflating the numerator.

So, we've had some discussion about this internally. I'd be interested in the steering committee's input actually about whether toe amputations should be removed from the specification.

MS. OWENS: So, Patrick, in our submission we submitted it with toe amputation excluded.

MR. ROMANO: Oh you did, okay.

MR. FRANCE: It's a bit confusing actually because under the ICD-9 codes listed here it says toe amputation is in the numerator. And then under the excluded case it says with any listed procedure codes for toe amputation.

MS. OWENS: Okay, so what that says to me is I need to go back to our documentation as well as our SAS code because clearly there's an inconsistency. Thank you for pointing that out.

So, back to Patrick's point then, does the steering committee want to have a conversation about it. And I apologize that it's neither clear to me nor in our documentation.

MR. ROMANO: Right. I think the intent was to remove it. But it appears that the language is not consistent. So, the intent was to remove it.

DR. SAMPSEL: So, if there is an interest in discussion on changing the specifications of this before consideration then we would want to put this measure on hold and not vote on it right now.

So I look to the committee on if you want to consider the toe amputation issue, if those adjustments need to be made. And

if they do then we would put this on hold and bring it back versus go through a vote on it.

MS. OWENS: So, if it's okay with you and it's possible to put it on hold, up to the committee.

But I would like clarification on what we're actually doing. I'm not sure that I can have a discussion when I can't articulate without seeing the SAS code what these numbers are based on. And I think you guys deserve to know that.

MS. MUNTHALI: I think Sarah was speaking in the interest of time. We're also trying to accommodate NCQA who has a hard stop in about 45 minutes.

And so we're saying put it in hold as in moving it a little further in the agenda. And I think there are a lot of issues we need to talk about. And we wouldn't be able to vote it sounds like to accommodate them as well. So I hope that's

fair. And I just responded to your email.

MR. SALIVE: Do you mean today or some other day? Because I think it affects -- I mean I'm just going to speak in favor of holding it, but I have other issues too.

MS. MUNTHALI: Today.

MR. SALIVE: That haven't been discussed.

MS. MUNTHALI: Today. This is just scheduling. It will be the next measure we talk about after the breast cancer screening measure.

DR. SAMPSEL: So, what we're going to do then is go ahead and put this on hold. We will come back to it after the breast cancer screening measure.

I know, Pam, you need to leave, correct? But thank you. And Patrick, are you able to stay on?

MR. ROMANO: Can you tell me when to come back on?

DR. SAMPSEL: Yes, we can send you

an email and let you know when to come back on since there are technically two additional PQIs in addition to this one.

MR. ROMANO: Okay, thank you.

DR. SAMPSEL: Okay. So we'll go ahead and move to breast cancer screening and at least the vote on this and additional discussion on this one on hold.

MS. OWENS: And for the committee's purposes AHRQ is still very much hearing what you have to say. Carol Stocks who just joined the AHRQ QI team will be -- is on the phone and will be listening to comments when the PQIs come back up as well as I will get a very detailed transcript.

So please continue to provide those comments because I need them to improve the measures as well as integrate them in terms of how we disseminate and what we put in our documentation. Thank you.

DR. SAMPSEL: So we're now switching gears to breast cancer screening and are

joined at the table with Sepheen Byron from NCQA. And we'll just get those slides up and get started.

Okay, so this is measure number 2372. We're trying to find the page in the workbook and we'll let you know. But we will have Sepheen do a brief introduction. And then Katie, are you prepared to discuss? Great.

We're also joined by Mary Barton.

MS. BYRON: Thanks, Sarah.

DR. SAMPSEL: Page 50. Sorry, Sepheen.

MS. BYRON: All right. So, this is switching gears significantly from all the PQIs that you guys have been discussing. But this is a health plan-level measure that looks at breast cancer screening in women 50 to 74 years of age.

Just some historical background here. This was a measure that was endorsed by NQF but it lost endorsement when it

became out of alignment with the U.S. Preventive Services Task Force guideline.

Previously the task force had recommended screening in 40- to 69-year-olds and that was what our measure originally specified.

And we have since updated the measure to align with that 50- to 74-year age range that the task force recommends. And so the measure has been updated and is now in the HEDIS health plan measure set as a measure specifying women 50 to 74 years for biennial, so every 2 year mammogram screening. So we're coming back for re-endorsement.

DR. SAMPSEL: Thank you, Sepheen. And I'll turn it over to Katie to start the discussion on criterion 1 and evidence.

MS. SELLERS: Okay, great. Yes, so this is a process measure. And the evidence provided is -- it's based on the U.S. Preventive Services Task Force

recommendation. It is a level B recommendation as are most cancer screening, or maybe all cancer screening recommendations.

It was -- the task force rating for the quality of the evidence was fair which is acceptable. It was based on seven randomized controlled trials. I think that's about all to say about it, about the evidence.

DR. SAMPSEL: Comments? Questions? Considerations about the evidence to support this measure? Eric.

MR. FRANCE: Just maybe to point out that it is being reviewed by the USPSTF now and so we'll be maybe thinking about the evidence based on their 2009 review which as you said was a B rating. Moderate certainty of moderate benefit.

And they're in the process now of redoing this so either our task force or a future one might want to rethink it once

they come up with new recommendations should they show less benefit associated with the evidence.

DR. SAMPSEL: John?

MR. AUERBACH: I would just ask how much confusion there is in the clinical community about this. Because my impression is there is quite a bit of confusion about what the appropriate recommendations are.

And so I ask that just because I think that using this as a measure of quality in a period of time when the recommendations are changing and that there is a good deal of uncertainty about what appropriate care is I just think is confusing for what it means.

DR. SAMPSEL: And I think I'd ask Sepheen and Mary if you can comment. Probably considerations of the MAP.

MS. BYRON: Yes. And I would say that one of the reasons that supports having the measure. You know, to be clear, the

measure does not penalize you for doing screenings in other women.

So you know, the task force also has a C recommendation which says it's an individual decision between a woman and their physician based on their individual factors for the lower age groups.

But our measure just says that where we do have evidence that is clear for the 50- to 74-year-olds that's where we would like to see screening happen.

MS. BARTON: Just that I would say that Sepheen and the team looked not only at the U.S. Preventive Services Task Force recommendation but at other clinical recommendations. And where they used the intersection set really where the recommendations overlapped to make the measure.

The measure is never going to replace a guideline and it's not our intent that people should act as though this

measure is a guideline. But rather as an indicator of where there is a consistent message from all of the recommending bodies it is in this age group.

MS. BYRON: And I would just add one more thing which is that the measurement advisory panel that we did convene, we always strive to keep them as multi-stakeholder as possible.

So we had representation from clinicians and oncologists but also patient advocates, policymakers, women's health, general internists as well to make sure that we got that balance. And that has helped us to develop the measure that we did.

DR. SAMPSEL: Marcel?

MR. SALIVE: So, I think even if the task force is reconsidering that it's not going to change its recommendation in this age group. So you know, that's a highly unlikely outcome. There's not a lot of new evidence being generated in that age group.

And so that's what it would take. So, they may change it in other age groups but this age group is fine and I think solid and can be endorsed probably.

MR. FRANCE: I'd just add that that may not be the case. They may be moving more towards a shared decision approach for this kind of screening.

It's -- breast cancer screening in women is -- when you look at the actual numbers it looks like prostate cancer screening in men in terms of lives saved and values. Every two years.

Finally, I'd just clarify too that cervical cancer is an A rating screening and colorectal screening is an A rating screening. I think the comment was made that most of them would be --

DR. SAMPSEL: Eric, we can't hear you. Can you speak?

MR. FRANCE: All right, I'll speak a little closer. Just to clarify that the

colorectal screening and cervical are both A-rated screening tests rather than B-rated as I think Katie had mentioned. And prostate is a D, against. So I would not be surprised if USPSTF changes its recommendations.

I think still our basis for today's vote is on the measure as is with evidence as reviewed which is the B rating and that's how I'll be looking at it.

DR. SAMPSEL: Ron?

MR. BIALEK: Just a clarification. So this is a screening ever. Just one screening. The way it's specified is -- okay.

MS. BYRON: At least one mammogram. In the two-year period.

MR. BIALEK: In the two-year period.

MS. BYRON: Yes.

MR. BIALEK: No, I'm just looking at the measure.

MR. FRANCE: Yes, the measure is

every two years.

MR. BIALEK: Okay.

MR. FRANCE: And so it's looking back in the last two years whether you were screened.

MR. BIALEK: Which is consistent with the task force.

DR. SAMPSEL: Okay. If there are no other comments or questions about evidence we'll vote on evidence and 1(a).

MS. ROBINSON-ECTOR: Voting is open. All votes are in and voting is closed. For evidence 5 voted high, 12 voted moderate, 2 voted low, 1 voted insufficient and zero voted insufficient with exception.

DR. SAMPSEL: Great. Katie, were there any comments regarding performance gap?

MS. SELLERS: The performance gap was pretty clear. The data -- well, what I wanted to say was that the data that were presented were all based on the formerly

recommended age group. So they were based on women 40 to 74 years.

But given that the performance gap was pretty clear. Looking at the different plans the means ranged from 50 to 71 percent. And some of the individual plans were really quite low. It looked like one was even 4 percent. So, very big performance gap.

DR. SAMPSEL: Questions or comments about performance gap? We'll go to vote.

MS. ROBINSON-ECTOR: Voting is open. All votes are in and voting is now closed. For performance gap 13 voted high, 7 voted moderate, zero voted low and zero voted insufficient.

DR. SAMPSEL: And Katie, any comments on priority?

MS. SELLERS: Yes, so as far as the priority goes I think there was consensus that this was a high-priority health condition.

About the screening itself I think that might be more of a moderate rating as, you know, a little bit of controversy over the benefit of the screening. But for the most part I think the priority was moderate to high.

DR. SAMPSEL: Discussion on priority.

MS. ROBINSON-ECTOR: Voting is open. All the votes are in and voting is now closed. For high priority nine voted high, nine voted moderate, two voted low and zero voted insufficient.

DR. SAMPSEL: And Katie, if you could lead us on reliability, validity, scientific acceptability.

MS. SELLERS: Yes. So, the numerator statement is women who received a mammogram to screen for breast cancer. The denominator is women aged 52 to 74 years as of December 31 during the measurement year.

And then the time period for this is

27 months. So that's the two years plus an additional three-month leeway which is used in other HEDIS measures in a similar fashion.

It has to be primary screening only, not biopsies, ultrasounds or MRIs. The data source is administrative claims electronic clinical data. There was some confusion in the workgroup around the time period but I think it's pretty clearly a two-year time period with an additional three-month sort of leeway which is to help with pushing for over-screening, just to make sure it doesn't push for over-screening.

And then on the reliability it seemed to have very high reliability measure scores. They did a beta binomial test. It was -- the values were 0.95 to 0.99 with a very large sample. It had 1,000 different plans representing over 80,000 patients.

DR. SAMPSEL: Questions about reliability and validity? Okay.

MS. ROBINSON-ECTOR: Voting is open. All votes are in and voting is closed. For reliability 12 voted high, 8 voted moderate, zero voted low and zero voted insufficient.

MS. SELLERS: Okay, so moving onto validity. The NCQA uses a systematic process for face validity testing which was shown to be strong.

For construct validity what they did was look at the correlation with colorectal screening and with cervical cancer screening at the plan level. And those correlations were 0.7 and 0.73.

They did not do risk adjustment as far as I could tell. The exclusions were for bilateral mastectomy. There was not an issue of missing data.

I guess one thing that I would note that I think was sort of alluded to earlier was that it does not allow for exclusion due to patient or provider refusal. But overall I think it looked like pretty high validity.

DR. SAMPSEL: Comments or questions? Go ahead.

MS. ROBINSON-ECTOR: Voting is open. All votes are in and voting is now closed. For validity 9 voted high, 10 voted moderate, 1 voted low and zero voted insufficient.

MS. SELLERS: So for feasibility these are data that are already being collected and the workgroup had no concerns on this.

DR. SAMPSEL: Committee members? Okay.

MS. ROBINSON-ECTOR: This is for feasibility and voting is open. All votes are in and voting is now closed. For feasibility 19 voted high, 1 voted moderate, zero voted low and zero voted insufficient.

MS. SELLERS: So for usability and use this is similar to a number of other measures we've seen where there's a lot of current use, it's used in health plan report

cards, State of Healthcare Report, Medicaid Adult Core set, NCQA accreditation, Quality Compass, et cetera, et cetera.

But the performance has been steady. So I think questions about usability have to do with has it been used for actual improvement.

As far as unintended consequences I think there's a potential for over-screening and the consequences that go with that. But again, that's with -- the three-month leeway is intended to help with that so that there isn't additional screening just to meet the measure. I think that was about it for usability and use.

DR. SAMPSEL: Ron?

MR. BIALEK: When the measure is published will it be specified exactly as it is on the screen without any period of time?

So right now it reads the percentage of women 50 to 74 years of age who had a mammogram to screen for breast cancer.

DR. SAMPSEL: Are you talking about publishing in HEDIS?

MR. BIALEK: When this measure is published by -- endorsed by us, by NQF is it exactly going to appear that way without any indication of the period of time.

So for instance, with oral health it says screening within the reporting year. This measure doesn't say anything about how often. I know it is in the data and it's taken into consideration, but the measure that we're endorsing would read exactly this way? Without any --

MS. MUNTHALI: Yes, it would read exactly this way. Unless the committee would like to, again, recommend to NCQA to make that refinement.

MR. BIALEK: My concern is that when John earlier said about confusion is this measure articulated this way going to add to confusion or help with confusion? Unless one digs into what the measure says it does.

MS. BARTON: We're sympathetic to the fact that there's clinical confusion. We specify this for health plans and in fact there's a volume that health plans get that includes the detail of the specification for how it's to be implemented.

I think if NQF is interested in a different level of detail for this particular four-line description which admittedly is super high-level -- there's a lot of details that are not included there -- I would be delighted to work with them and make sure that you have all the relevant data points you want to put into your four-line summary.

MS. MUNTHALI: And this is just for presentation for today. It's just a really brief synopsis of what's in your submission form. So I think we need to -- maybe the two of us can work together. You could reevaluate your submission form to make sure that that's comprehensive enough and it's

not causing confusion when we put it out.

DR. SAMPSEL: Jacki.

MS. MOLINE: There is so much going on now. I think now the breast surgeons are beginning to weigh in. They've convened a panel to decide what the recommendations should be because there is such controversy.

But in many ways it shows the value of the measure which is to say where are we now. And how over time with the changes in the professional organizations or the changes in the USPSTF, how do the rates of mammography change in accordance with different guidelines.

So in fact, if anything it makes the measure more valuable because it's a way -- it's almost a snapshot of getting a sense of how do patterns change with controversy and/or deferential opinions from different professional organizations.

And it's actually -- it's a social experiment in many ways to look at this. So

there's a way of looking at the measure in that way.

DR. SAMPSEL: Mike?

MR. STOTO: I think that Ron has an important point. And if we just added "in the previous two years" to the end of that it would simplify things quite a lot.

People tend to grab this and put it on a figure and so on and that's an important thing to do also.

DR. SAMPSEL: And I'll just add with my health plan hat, you know, during the time period that this measure had lost endorsement certainly health plans didn't take priority off the measure. But it was -- there was an area of confusion of what do we work on. And how do we move forward with some of our quality improvement efforts.

So I think, you know, your point earlier was valid, Katie, about kind of questions on where are there QI efforts right now.

And you know, we saw and I was working for WellPoint at the time, so we also saw different interpretations across the country on which age groups to use. So people have been waiting for this measure to come back.

Not that it really ever went away, but having some more of that standardization around for comparability because folks were starting to use just iterations of unendorsed measures which didn't make the field happy either. So I do think it's still in heavy use and of strong interest to health plans.

So with no other comments we'll vote on usability. Oh, I'm sorry, Renee?

MS. FRAZIER: I just want to clarify. Are we going to add the comment or not? I mean, that's what I was going to -- and Ron said it. Are we or aren't we adding that simple language in this summary format so it's clear and up front what the endorsed

measure is?

DR. SAMPSEL: Right. So the specification already clearly says that. What we'll need to work on is on the description for the NQF quality position system website.

MS. ROBINSON-ECTOR: This is for usability and voting is open. All votes are in and voting is now closed. For usability 14 voted high, 5 voted moderate, 1 voted low and zero voted insufficient information.

DR. SAMPSEL: Okay, other comments regarding overall suitability for endorsement? I think this still goes, you know, this is probably where that description conversation came back in. But any other concerns, questions? Great, let's vote.

MS. ROBINSON-ECTOR: For overall suitability for endorsement voting is open. All votes are in and voting is now closed. For overall suitability for endorsement for

measure 2372 Breast Cancer Screening 18 voted yes, 2 voted no. The measure passes.

DR. SAMPSEL: Great. Thanks, everybody, for making it through the morning. And lunch is here so we're going to ask that everybody get up and grab their lunches.

We're going to reconvene at 12:30 and I've asked Patrick to call back in at that point to continue our discussion on the PQI measures.

(Whereupon, the foregoing matter went off the record at 12:11 p.m. and went back on the record at 12:36 p.m.)

DR. MCINERNEY: We've had hopefully ample time for most of you to get most of your lunch. And my understanding is that we postponed a discussion on 0285 which is the rate of lower extremity amputation among patients with diabetes.

Oh, I'm sorry, public comment first.

MS. MUNTHALI: Cathy, can you please

open up the lines for members of the public to make comment if they'd like?

OPERATOR: Yes, ma'am. At this time to make a comment please press * then the number 1. There are no public comments at this time.

MS. MUNTHALI: Thank you.

DR. MCINERNEY: Okay, so this is on page 24 of your worksheet. Did we do any discussion of this at all?

MS. MCKANE: We began to have some discussion and we went through the evidence.

DR. SAMPSEL: Yes, we went through evidence and then we were having a discussion regarding --

MS. MCKANE: Regarding whether toe amputations are currently excluded from the numerator. But they're included in the Excel spreadsheet as one of the ICD-9 codes that's included in the numerator. So we needed to have clarification on what the measure was actually -- what the actual

measure was.

DR. MCINERNEY: Are our measure developers here?

MS. MUNTHALI: Patrick, are you on the line?

DR. MCINERNEY: Are we expecting them to join us?

MS. MUNTHALI: We are. He was coming from a conference. I told him what time but I don't think he's out of the meeting yet.

DR. MCINERNEY: And what about the other two measures that we haven't --

MS. MUNTHALI: They're also AHRQ.

DR. MCINERNEY: AHRQ. Oh, well.

MS. MUNTHALI: I think let's just break for about 10 minutes and we'll try and reach Patrick some other way. And then if not we can proceed.

DR. MCINERNEY: All right, so should we just not do anything?

MS. MUNTHALI: For 10 minutes.

DR. MCINERNEY: All right, enjoy your lunch and take a little postprandial nap if you'd like.

(Whereupon, the foregoing matter went off the record at 12:39 p.m. and went back on the record at 12:49 p.m.)

DR. MCINERNEY: Patrick is with us, folks, so we're going to start on measure 0285 on page 24 of your measure worksheet. This is the rate of lower extremity amputation among patients with diabetes.

And the description of the measure is admissions for any listed diagnosis of diabetes and any listed procedure of lower extremity amputation per 100,000 population ages 18 and older. And it excludes any listed diagnoses of traumatic lower extremity amputation admissions, toe amputation admission likely to be traumatic, obstetric admissions and transfers from other institutions.

And I understand that there was some

previous conversation about this this morning and some questions were raised. And could we review those questions and then proceed, please?

MS. MCKANE: Sure. This is an outcome measure and we were discussing the discrepancy between the numerator as listed in the description versus the numerator that's described in the ICD-9 codes. Because the ICD-9 codes indicate that toe amputations would be included if within the numerator. And the description of the numerator excludes toe amputations. So we wanted to have some clarification on that issue from the developer.

DR. MCINERNEY: Patrick?

MR. ROMANO: Yes, can you hear me?

DR. MCINERNEY: Yes.

MR. ROMANO: Yes. So we were able to clarify that. And so, this is basically an idiosyncracy in the SAS code. Suffice it to say that toe amputations are excluded.

The brief historical perspective is that in the original version of this indicator toe amputations were included. And subsequently based on user feedback as well as empirical analysis and a second round of expert panel discussion.

We did choose to remove toe amputations in version 4.5 which is the current version. But the way that was done was basically by backing them out after putting them in.

So it's a little bit confusing with the technical description but the intent as well as the operationalization is that they are excluded. So if there's only a toe amputation with nothing more proximal the intent is to exclude. And that is what the SAS code does.

DR. SAMPSEL: Thank you.

DR. MCINERNEY: All right, is everybody comfortable with that? Yes? Okay.

So who is going to lead the discussion?

MS. MCKANE: I'm leading the discussion.

DR. MCINERNEY: Oh good.

MS. MCKANE: And we -- I'm not really sure where we left off but we did talk about that there was evidence. This is like the other indicators where it's an outcome but there are other factors beyond -- in access to medical care, or medical care that could be in place such as community-level factors. So it was very similar to the other measures in this family.

And I believe that we identified there is a performance gap. There is evidence that the rate as described in the description from the developers is increasing yet that rate is not adjusted for prevalence of diabetes and is not -- is per 100,000 of population and not among diabetics. And that is something that was a

concern that was raised with this measure.

And when we look at out accounting for diabetics -- among diabetics the incidence of amputations is actually decreasing.

We did note that there were gender differences and age differences. Race and ethnicity is not mentioned as something that was measured but was based on the literature. But there is definitely racial and ethnic disparities.

We felt the priority was high. It's a high-impact priority for a smaller segment of the population. But the prevalence of diabetes is increasing. And the measure may capture people who had little interaction with healthcare prior.

We also felt that, let's see, I don't -- we also had made some other comments with regard to using geocoding to give them a more precise estimate validity to pinpoint high-risk neighborhoods.

So I think that that is a summary from the workgroup. Are there any more comments about that?

I know that we had quite a discussion about all of this in the call as well as around these other measures as well.

DR. MCINERNEY: Other comments from the committee members on the evidence? All right. Oh, yes.

MR. SALIVE: So I did look into the business of why the rate didn't change. I don't know exactly if we ironed that out.

But when I looked at the national data from CDC on where it's -- it is a Healthy People 2010 objective to lower the risk of amputations. But they do use a denominator of diabetic patients.

And they did report that it was cut in half from 1997 to 2005. So, there was -- that was another source of information that I found.

And I wasn't sure if this was the

time but I do think harmonization with that national objective would be helpful. By you know, using the diabetic population as the denominator. But maybe that's for now, or maybe that's for later.

DR. MCINERNEY: I think that's a good point about the harmonization. And I think though we would save that for later. Are we ready then to -- oh, sorry. Mike?

MR. STOTO: Is having the denominator only be people with diabetes more important for this one than for the other diabetes measures? I'm wondering why this is coming up now? Why is that?

DR. MCINERNEY: That's a good question.

MR. SALIVE: There are pros and cons to using the diabetes persons as the denominator. And so Census data is very widely available. The prevalence of diabetes does seem to change partly because a lot of it is undiagnosed. And so I think

there is a trend towards greater diagnosis of diabetes now than in the past.

As well as a change in strictly speaking the amount of diabetes. So, I think getting -- so getting small area diabetes estimates is the tricky part.

MR. STOTO: But that -- everything you said also applies to all the diabetes measures that we've been talking about. Presumably. I don't understand why this one is different from the other three with respect to what the denominator should be.

MR. SALIVE: I feel strongly that it should be for this one.

MS. MCKANE: I think it was something we talked about in the call if I remember because we also -- for all the measures, particularly the longer-term ones.

But I think that one of the takeaways that we were talking about is the importance of having a population level, the population, and then also trying to

attribute it to the -- among the cohort that actually has diabetes.

And the tricky part is actually trying to define the number of patients that have -- or to reach a denominator that is meaningful for patients with diabetes was the issue that I think we kind of bumped up against.

And it is something I thought in the specs that these could be analyzed in both ways, that that was not -- for any of them, for any of the measures.

DR. MCINERNEY: So I think for consistency's sake we should stick with what we've been doing and that is for the entire population.

MR. STOTO: I think that having one with a different denominator than the rest would be far more confusing.

DR. MCINERNEY: Right. Good. Okay, any further discussion on evidence? All right. Let's vote on evidence, please.

MR. ROMANO: This is Patrick. Could I say something?

DR. MCINERNEY: Sure, Patrick.

MR. ROMANO: Yes, just two quick comments. One is that AHRQ has been doing empirical work over the last year to try to implement a small area estimation procedure for diabetes prevalence using the BRFSS data, the Behavioral Risk Factor Surveillance System data that are available from CDC.

The practical issue is that we need estimates at the county level and we need estimates also that can be drilled down to age and gender subgroups within small areas.

So, anyway, the empirical work is ongoing. And so we may come back to NQF, AHRQ may come back to NQF with a recommended implementation procedure. But it's not technically straightforward because of the desire to be able to offer these rates, not just at the state level where it would be

easy, but at the county level or the MSA level where it gets trickier.

MR. AUERBACH: I would just say that the dilemma in terms of the database that Patrick's talking about, it's self-reported random digit-dial telephone calls asking people if they're diabetic. And so it's a really inexact measure.

So it's often what gets used but I think it doesn't account for undercounting or under-diagnosis. And it also doesn't count -- it's not done in multiple languages. There's a lot of obstacles to using that as a reliable data source.

So I think that just having a denominator of diabetics is quite challenging.

MR. ROMANO: Yes. And the other point I wanted to make is with reference to the time range, just so we're talking about parallel time ranges.

Actually, if you take the AHRQ

indicator back to 2000 there has been a significant drop during the period from 2000 to 2008. It's just that the rates have been flat since 2008. So between 2000 and 2008 there was approximately a one-third drop even according to the AHRQ indicator in lower extremity amputations for diabetes per 100,000 persons.

MR. BIALEK: Having absolutely no knowledge or expertise in this topic area I wanted to ask a question of those who do.

Is there anything that's lost by these changes that are being made in terms of -- we're capturing events, right? Not people. And is there any problem with that? Any concern about just capturing events, not people.

Removing the toe amputation sounded like that reduces that problem but does it reduce it enough? Because I think we want to be capturing people versus just events.

DR. MCINERNEY: I mean you're

concerned that the same person could come in for two or three different amputations at two or three different levels.

(Laughter)

MR. BIALEK: Well, and there's parts of -- I mean, I just don't know enough to -- because to make a change we're talking about changes to individuals, right?

So is it a problem? It may not be. I just am asking the question.

DR. MCINERNEY: Patrick, can you answer that, please?

MR. ROMANO: Well, it is still potentially an issue.

Of course, one might argue that to the extent to which the left leg has amputation and the right leg later, that those perhaps should count as separate events. So it depends whether the right unit of analysis is the person or the extremity. And we haven't resolved that question.

But we do think that by removing the toes it removes the most obvious manifestation of this problem of non-independence.

DR. MCINERNEY: Thank you. Are we ready to take the vote then on the evidence? Importance to measure evidence? All right, please, Kaitlynn.

MS. ROBINSON-ECTOR: This is for evidence and voting is open. I think we're waiting for one more vote. All votes are in and voting is now closed.

For evidence 18 voted yes and zero voted no.

DR. MCINERNEY: Okay, good. So then I think we also -- we did have some discussion about the performance gap in the priority in your presentation.

Do we have any further discussion on performance gap or are we ready to vote on that now?

It looks like we're ready to vote on

performance gap, please.

MS. ROBINSON-ECTOR: This is for performance gap and voting is open. All votes are in and voting is now closed. For performance gap 13 voted high, 5 voted moderate, zero voted low and zero voted insufficient.

DR. MCINERNEY: Okay, moving along now, priority. High-priority. Any further discussion on priority?

Okay, then let's vote, please. Thanks.

MS. ROBINSON-ECTOR: This is for high priority and voting is open. All votes are in and voting is now closed. For high priority 15 voted high, 2 voted moderate, 1 voted low and zero voted insufficient.

DR. MCINERNEY: Okay, good. Let's move now to discussion of reliability and validity, please.

MS. MCKANE: The numerator and denominator are well defined. The

denominator is from the Census data. And the data source for the numerator is the discharge data that we've been talking about.

The measure was tested for reliability at the measure score level. And overall the measure was rated as moderately reliable by the developers. And it seemed to be it was more reliable for larger population areas and very much consistent with the other ones.

As far as -- are we talking about validity too? So, the specifications align with the evidence. The measure was tested for validity at the measure score level. And they did empirical validity testing. And I think we rated this as moderate or moderately high.

We did have some -- there was a comment from -- that was submitted from the public that the -- from the Armstrong Institute for Patient Safety and Quality who

questioned whether this was the correct denominator which we were discussing. But also made the comment both the numerator and the denominator are easy to collect.

MR. AUERBACH: Are we talking about reliability and validity at this point, or just reliability?

Well, I guess -- this is in the spirit of -- for who's often asked this question. I guess I just wanted to make sure that we are somehow going to capture in the commentary on this the confusion one could have in looking at hospitalizations and rates appearing to go up while we think that they've actually gone down if we have the denominator, an accurate denominator with diabetes.

Because again, if you're simply looking at that and you're using it for quality standards it looks like you're failing because the numbers are going up. And so I think it just requires some

crafting of language to help people to interpret what it means.

If the story is we're doing a really good job even though the number is going up and that insurers and providers should be aware of that I think crafting that language is important.

MR. VALDEZ: Could the developer just explain to me again, I'm not sure I understood why transfers from other facilities were being excluded from this measure?

MR. ROMANO: That would be to avoid double-counting essentially of the same hospitalization. When the hospitalization -- into two different facilities.

MR. VALDEZ: But you're not looking at the hospitalization, you're looking at the amputation.

MR. ROMANO: Right. So, your point is that --

MR. VALDEZ: -- take the leg off

twice.

MR. ROMANO: Yes.

MR. AUERBACH: It's the question of whether the measure of quality is the ambulatory care or the nursing home care I think that determines it.

MR. VALDEZ: Right, but we're not looking at care necessarily. Otherwise we'd go to the diabetic as the denominator.

If we're looking at amputations then it shouldn't really matter where the source of the patient is.

MS. MCKANE: Right, and I had the same question for the developer about the exclusion of nursing home patients. And the rationale provided was that they are -- they should have -- be surrounded by care.

And the reality is that may or may not be true and for the same reason, that part of the population within that area is another amputation that's not being counted.

MR. ROMANO: Yes. Well, I think

that the point, certainly we accept the point reference to transfer from another hospital. It's quite possible that that exclusion should be revisited.

Because as you're saying if we're only counting the hospitalization at which the amputation occurred there isn't any clear reason to exclude patients based on whether they were admitted to another hospital first.

In the case of skilled nursing care that is an across-the-board exclusion because these indicators have been viewed as ambulatory care-sensitive conditions or indicators. And by definition if somebody is a long-term care resident of a skilled nursing facility they're not receiving ambulatory care through the same healthcare structures as someone who's in the community.

But, so those two issues are slightly different and I think the first one

certainly should be reevaluated. And we can put that on our to-do list to check what the implications of dropping that exclusion would be.

The second related to SNF is a broader conceptual issue that I think underlies all of these indicators. So, I'm not sure I see that changing without much broader discussion.

MS. MCKANE: Thank you. I was just wondering if in the future are you planning to measure this for the elderly population? Or do you actually have a measure?

MR. ROMANO: There is no AHRQ measure specifically for the long-term care population. I think CMS obviously has other measure development programs that are geared towards the long-term care population through the Oasis data. But I don't think there's an indicator of this type in the CMS set. But certainly that's a topic that could be brought into harmonization

discussions with CMS.

DR. MCINERNEY: Okay. I think we're ready to vote on reliability, please. Thank you.

MS. ROBINSON-ECTOR: This is for reliability and voting is open. All votes are in and voting is now closed. For reliability 6 voted high, 12 voted moderate, zero voted low and zero voted insufficient.

DR. MCINERNEY: Validity. Do we have any further discussion about validity? All right, hearing none let's go ahead and vote on validity, please.

MS. ROBINSON-ECTOR: This is for validity and voting is open. One vote is still out. All votes are in and voting is now closed. For validity 5 voted high, 11 voted moderate, 2 voted low and zero voted insufficient.

DR. MCINERNEY: Some discussion on feasibility, please.

MR. ROMANO: Can I ask a

clarification question of NQF staff?

So, if we were to recommend implementation, for example, of dropping the hospital transfer exclusion based on the discussion here what would be the process for doing that? Would that have to come back to this committee, or would that be done at a staff level?

MS. MUNTHALI: It would come back to the committee and we'd ask the committee to re-look at it after comment. And then vote over a phone call.

MR. ROMANO: Okay, thank you.

MR. VENKATESH: Patrick, this is Arjun. If we're going to look at that I would only ask if it's possible to do an analysis where you look at hospitalizations in the previous or following day from the hospitalizations that are being currently excluded because of the transfer.

Because we've done this for vascular surgeries before and find that often there

are two hospitalizations with the principal discharge diagnosis of the vascular surgery even though the procedure was only done during one of those hospitalizations.

So we need to check like that to make sure we're not creating some unintended counts.

MR. ROMANO: Thank you.

MS. MCKANE: Okay, move onto feasibility. This is very similar to all the others. This uses electronic hospital claims and Census population denominator.

The data are generated during care and coded. And it seems to be the data collection are implemented --

DR. SAMPSEL: Patricia, can you speak closer to your microphone?

MS. MCKANE: Oh, I'm sorry. The feasibility we felt was -- we felt was pretty high.

We did have concerns about the fact that race and ethnicity data aren't

collected in this data source. And we do know there are disparities. But other than that that was -- electronic medical claims data.

DR. MCINERNEY: Further discussion of feasibility? Let's vote, please.

MS. ROBINSON-ECTOR: This is for feasibility and voting is open. All votes are in and voting is now closed. For feasibility 13 voted high, 5 voted moderate, zero voted low and zero voted insufficient.

DR. MCINERNEY: Usability, please.

MS. MCKANE: It's widely used by DHHS, AHRQ, CMS and numerous states. So we don't really have many additional comments regarding the usability.

DR. MCINERNEY: Any further discussion? Okay, let's vote on usability, please.

MS. ROBINSON-ECTOR: This is for usability and voting is open. There's one vote out. We're still missing that vote.

All votes are in and voting is now closed.

For usability 14 voted high, 4 voted moderate, zero voted low and zero voted insufficient information.

DR. MCINERNEY: Okay, we're ready for a discussion on overall suitability for endorsement. Any further discussion? All right, let's vote, please.

MS. ROBINSON-ECTOR: This is for overall suitability for endorsement and voting is open. One vote is missing. All votes are in and voting is now closed.

For overall suitability for endorsement measure 0285 Rate of Lower-Extremity Amputation Among Patients with Diabetes (PQI 16) 15 voted yes, 3 voted no. The measure passes.

DR. MCINERNEY: Very good. Down to two more.

Now, the discussion I believe this morning was to do 0280, dehydration first. All right. That's on page 14. And the

description of this measure, it's admissions with a principal diagnosis of dehydration per 100,000 population ages 18 years and older. And it excludes obstetric admissions and transfers from other institutions.

Let's see, who is going to lead this discussion?

MR. VENKATESH: That's me. Okay. So this is very similar to all the other measures with respect to its denominator. The difference here is it's trying to look at inpatient admissions for what is -- while the title says dehydration it's really composed of a set of codes associated with dehydration as well as adult gastroenteritis.

And so some of this measure overlaps with yesterday's pediatric measure on gastroenteritis but not entirely. Just think of it as kind of including that as well.

And so in general I guess, and I

mentioned this to Pam when she was here before, my view is that this measure may be one that we should think about retirement in. Because since this measure was developed in the mid-two thousands the healthcare system has evolved.

And I'm not sure this is either really measuring what we want to be measuring and what it is measuring may not be particularly helpful or useful.

And so the background I'll give to that is that over the course of the last 506 years dehydration has changed for a couple of years. One is that the coding of it has simply changed. So even though the same care processes might be occurring a lot more happens in observation services. It's the same hospital bed, same everything. The only difference is that it's billed as an observation visit.

The other thing that's changed is also the care has changed and become more

ambulatory. People are able to get IV fluids in the ambulatory setting. It's moved more towards the emergency department or an actual observation setting.

And so hospitalization for dehydration as a whole is down and reflects a set of people that may not necessarily be the preventable dehydration. So with that background I'll kind of go through I guess each section.

For evidence this is an outcome measure similar to the others. I think it's important to think about whether there's a process-outcome linkage. And in this case the original application notes that support for fluid intake by high-risk individuals, those are people with cognitive or psychiatric needs, older age, comorbid illness, high-risk medications, could with additional support potentially have a preventable hospitalization.

And then they also said that a

community-level process that could potentially improve this is air conditioners during the summer heat.

The challenge is that there's not a lot of evidence base that suggests that these various processes actually reduce the hospitalization. But I actually think there's some face validity, a feeling that some dehydration hospitalizations are preventable.

They cite four clinical guidelines. The clinical guidelines that are cited are all level 4 or level 5 evidence, so non-experimental studies or expert consensus that are largely about how to -- from the nursing guidelines on how to manage dehydration in the geriatric population.

Those reflect possible processes that could improve dehydration care, but again, they're not linked to the outcome. So to me the clinical guidelines shouldn't really affect how we rate the evidence here.

So, ultimately where I left kind of that initial evidence measure was kind of as moderate. And that's largely based on the fact that it seems like there's probably some face belief to the fact that many dehydration hospitalizations can be moved to the outpatient setting.

That said, if we're very restrictive when we think about the way they've described the outcome which is inpatient hospitalizations for dehydration they may be much less preventable. These are people that likely have more comorbid illness and a variety of other things going on and so it may not be as applicable.

Should I go all the way through evidence or stop there? I think we stop there, is that right? Okay. So we'll do gap and opportunity for improvement later, right?

MR. FRANCE: Arjun, I'm curious about process here with the outcome as the

measure. Is it the pass/no-pass of the algorithm? Or is it down in the process where you're giving it sort of a low/moderate/high view?

MR. VENKATESH: No, it's not an outcome measure. So it's still in that top category.

The difference is that when you have an outcome measure we're also asked to think about whether or not there's evidence to support that processes can change the outcome. And I think there was evidence that processes can change emergency department use observation use for dehydration.

For inpatient hospitalizations there is evidence as well, some that's older, not really reflective I think as much of how care is delivered now.

DR. MCINERNEY: Patrick, would you like to comment on Arjun's presentation, please? Patrick?

MR. ROMANO: Yes, sorry. I think they're valid points. I'm not sure what Pam said when she was there in person, but certainly there has been a shift towards observation care and ED management of these patients with dehydration.

The question is it has perhaps changed the meaning of the indicator. Is it -- does that mean it's time to retire it? That's a little harder to say because it depends on your perspective on whether you think it's good to continue to encourage this move towards avoiding inpatient stays for patients with mild to moderate dehydration.

MR. CARILLO: Just a couple of points. Isn't the fact that hydration in an ambulatory setting may forestall or inhibit the admission, isn't that a measure of good quality care in the ambulatory sector?

And also, historically, for decades the EDs have hydrated you to try to prevent

an admission. So, I don't really see that there is a historical change that would make us go in the direction of dropping this.

MR. VENKATESH: So I think part of their historical change there is more driven by a change in payment policy than it is by massive change in care. There's certainly been a change in care where you can do much more rehydration in the ambulatory setting via other things.

But it's more payment policy. We've essentially in five years in both not just the Medicare population which is a lot of people are thinking about that role, but in the commercial population said that hospitalization for dehydration will be billed as observation.

And so what that means is in the data that we then look at, in the HCUP data that's used for the indicator or in anything that's based on inpatient claims you don't capture those.

And I'm not saying that's a bad thing. I'm with the Patrick in the sense that you could say, okay hey, measuring this as an inpatient hospitalization if we want to help continue to promote these going into observation then that's fine except that I don't think that this measure is going to drive people to move it to observation. I think that's purely based on payment policy. When an insurer or Medicare says hey, that's observation, then it's going to get billed as observation.

And so that's what's happened. That's why there's such a rapid change in five years, it's just the payment policy changed. And the payment policy made it unmeasurable in inpatient data.

MR. CARILLO: But the issue of observation applies to every measure in the book because that's happening across the board. So I don't see how that happening in this particular measure gives it any less

credence.

MR. VENKATESH: So I think it's that some of these diagnoses are more sensitive to observation than others. And so many of these, you know, some of the diabetes ones will still be frequently billed as a full inpatient stay. And so the degree of the change is less. It certainly applies to all of them.

I think you can see it in this one in the data because the rate has dropped by 40 percent in five years. And so I think everybody could when they see that say that it's really unlikely that the amount of dehydration at the county level dropped 40 percent in five years. These other measures are changing at much smaller numbers than that.

And so I think that's why the difference with this indicator, the other ones is the change in the payment policy has actually made the indicator not as reliable

or valid versus I think that the indicator carries more meaning for some of the others.

MR. CARILLO: But asthma, pediatric asthma, I mean that certainly has the same trend.

DR. SAMPSEL: I want to tease out a little bit from other members of the group to react to that and the concept of is this a measure that there may be or should be a consideration of a recommendation for retirement.

MR. SALIVE: So, I guess I would say that if it dropped 40 percent that's good, but I mean, you know, you're not the only payer. This is not the only, you know. There's still room for this to have some movement in the future. So I think it's a bit premature to recommend that. I don't see why we have to judge that.

I mean, it's like maybe it will go down to zero. But maybe something else will come up. I don't know. I think it's, you

know, it's a reasonable measure. There's a lot of measures. And I think -- I don't see a compelling reason to do it now.

MR. STOTO: I want to support Arjun on this one. From what I've heard you say it strikes me that the change that we see probably is due to changes in billing rather than changes in care, or risk, or anything having to do with the health of people.

And if we're seeing that over time we're probably seeing apparent differences across geographical areas that also aren't factual as well. And so if we have a measure that is picking up more changes in billing rather than changes in care that's not a good measure from what I can see.

DR. MCINERNEY: However, could one argue that the changes in billing have pretty much taken place by now and there won't be that much more change in billing from now on? And so that therefore this still could be a reasonable measure.

MR. STOTO: I have no idea.

MR. VENKATESH: I think that question then to the committee is if we think that's the case then is a measure of inpatient hospitalizations for dehydration something that we think is a good prevention quality indicator for a community.

Because I think the population of people who have inpatient hospitalizations for dehydration are probably different than some of the people we're thinking about in our head who could otherwise be managed in ambulatory settings, or who other forms of hydration could prevent hospitalization. It's a different pool of people.

DR. MCINERNEY: Okay. So I think we have some -- we've had some significant discussion. I think it's now time to vote and let's see -- oh, somebody else has something.

MR. FRANCE: Just a quick support of what Arjun just said. So, while Mike was

talking about the billing I think it's more compelling to talk about who's actually in the hospital today for a diagnosis of dehydration and are those cases that could have been prevented had they been taken care of differently or had they the health system infrastructure in place to manage them.

And Arjun's already mentioned that they're different. They're not the classic, I'm dehydrated, I go to the hospital. To prevent that as a monitoring in the county of how well we're doing in our system.

So I don't know that I have enough information to truly follow through with your description that these people are significantly different. So maybe that's the one piece that would be interesting to understand a little bit better.

MR. STOTO: To me this is actually a validity issue rather than an importance. I don't think the importance of dehydration has changed. It just strikes me from what

I've heard that this measure is not really picking up what actually is happening.

MR. CARILLO: The fact that there's oral rehydration and more IV treatment in the ambulatory does reflect good ambulatory care. So, I think that that's -- goes in line with what this measure is supposed to show.

MR. ROMANO: This is Patrick.

DR. MCINERNEY: Go ahead.

MR. ROMANO: There are data from some of the states participating in the HCUP program related to observation stays. And Pam may be able to comment further on that when she's back with us.

So, I'm not sure exactly if AHRQ has looked empirically at if there's a clear substitution effect between inpatient stays and observation stays. But we could potentially find some data on that question.

The other thing just in general context there are two PQIs that have not

been brought to NQF for endorsement precisely because they've dropped about 70 to 80 on or more because of changes in practice, specifically admissions for hypertension and admissions for angina without procedure.

Those two PQIs have not been brought to NQF for endorsement. They've been discussed for retirement because of the magnitude of the drop.

So, this kind of 40 percent drop is not yet of the magnitude where we would ordinarily recommend retirement. But we appreciate the discussion of the question.

DR. MCINERNEY: Thank you. Okay, I think we're ready to vote now on evidence. Please, Kaitlynn.

MS. ROBINSON-ECTOR: Voting is open. All the votes are in and voting is closed. For evidence 15 voted yes and 13 voted no.

DR. MCINERNEY: Three voted no.

MS. ROBINSON-ECTOR: Three, sorry.

DR. MCINERNEY: Okay, good. How about performance gap, please.

MR. VENKATESH: I would just say that -- so the main data here shows that there is still persistent variation between counties. But like I said, there was a 40 percent or so reduction between that five-year period.

And since we don't -- it's kind of the discussion Jason brought up earlier. You know, what's the right way. At what point have you improved? I don't know.

But my guess is that that's not that there was substantial improvement over that time and simply just billing change. And so at the end of the day I think I'm kind of left with, okay, there's variation between counties. I don't know how much improvement we've had or not in that time period.

And there is also -- sorry, there are disparities. And so actually I think at the end we kind of just left this at

moderate. Given that older adults are at higher risk there's been declines across all races but less so for Blacks as well as decline across all incomes but less so for lower-income areas.

DR. MCINERNEY: Thank you. Further discussion on that? Okay, let's vote on performance gap, please.

MS. ROBINSON-ECTOR: This is for performance gap and voting is open. Three votes are out. Three votes are still missing. One vote was missing.

So, 6 voted for high, 12 voted for moderate and 1 voted for insufficient.

DR. MCINERNEY: Okay. Some discussion about priority.

MR. VENKATESH: Dehydration is something that has fairly high frequency, carries some associated morbidity, a little bit more mortality, but it's also not necessarily a condition that's listed in the high-impact conditions of the HHS list or

National Quality Strategy and other major national priorities. So I kind of left it at something probably around moderate.

DR. MCINERNEY: Thank you. Any further? Mike? Okay, let's vote on priority, please.

MS. ROBINSON-ECTOR: Voting is open. All votes are in and voting is closed. For high priority 4 voted high, 14 voted moderate, 1 voted low.

DR. MCINERNEY: All right, thank you. Now we can get to the reliability and validity discussions, please.

MR. VENKATESH: So, the numerator of this measure is -- it was included above. I guess we can discuss it within validity. So, reliability.

There's two forms of reliability testing. They can either test data elements or the score similar to all other measures. The score has been what's tested.

In this case though I think that

data element reliability may be more -- is something that's also important. And the reason is that for a lot of the other PQIs, the diagnosis codes have fairly good fidelity to what's being measured. So an asthma code means it's asthma.

The challenge with dehydration is that it's vague and ambiguous and it's assigned at hospital discharge. So regardless of how long the patient was in the hospital, one day, three days, seven days, it's dehydration on the way out the door looking back. And so as a result it's much more difficult to say that that was the reason the person was hospitalized in the first place.

And there's codes for things alternatively you could think of that patients who are dehydrated would get hospitalized for. There's a set of codes around weakness and malaise, for example, that would often end up there.

And so, as a result, I'm not exactly sure of the universe of dehydration, how much this measure captures. They haven't done any chart validation to say what that would be. And so I think there's a little bit of a reliability concern that comes from that.

My guess is that this is just not capturing the universe. It's capturing a smaller subset.

And then otherwise I thought the reliability testing with respect to the score itself was really good. It's similar to what's been done before. The signal-to-noise ratio was also very similar. And kind of similar to other ones as well.

Larger areas seem to be probably measured with a little bit more reliability than smaller areas.

DR. MCINERNEY: Thank you, Arjun. Any other discussion? Yes.

MR. FRANCE: Just clarifying. With

principal diagnosis as dehydration do you still have that concern about discharge?

MR. VENKATESH: Yes. So principal diagnosis is not the admitting diagnosis, it's the discharge diagnosis.

And so it's fine if you think that the admitting diagnosis is pretty close to the discharge diagnosis which in the case of a lot of these things is going to be very close. Like amputation I'm sure is very tight. Things like asthma, very close.

In the case of dehydration it's being assigned by kind of the coder and the provider at the end on the way out the door. It may be very different than what originally happened kind of up front.

So you could see a patient with CHF who is on Lasix, over-diuresed and gets dehydrated and gets hospitalized. Are they hospitalized for dehydration or their CHF? On the way out the door the principal discharge diagnosis could be dehydration and

we would be capturing then something around their ambulatory care of their medications.

But more likely than not it's going to be a CHF-related diagnosis on the way out the door. And so those types of patients would not be captured.

Similarly, the other population that would not be captured are cancer patients who may frequently be admitted for dehydration but will leave the hospital with a discharge diagnosis around their cancer and therefore not captured by the measure.

MR. SALIVE: So, I guess I'm not so concerned about CHF being missed because I believe that is a different issue, being out of tune on their cardiac meds.

I appreciate the point but I do think that there are a number of other diagnoses listed in this paperwork about, you know, that do make your point a little bit better perhaps that viral gastroenteritis, they figure out exactly

which one it is and that's the principal diagnosis. But that's still included here.

A lot of kidney disease where they become dehydrated from their kidney disease. I think there's a pretty long list of diagnoses here. It's not purely dehydration codes. So, I'm pretty happy with this list and I have no concern.

MR. STOTO: I don't know anything about the subject other than I've heard just in the last few minutes. But it strikes me that this situation was quite likely that a patient in one location would be treated very differently from a patient in another location with respect to how the coding is done. And that -- which is a reliability problem for comparisons. It's probably true over time too. It sounds like there's a number of somewhat arbitrary decisions that have to be made but can be made in different ways in different hospitals and so on that may make it difficult to compare the results

that come out.

MS. MOLINE: With respect to the cancer that had been one of my initial concerns. But when you look at it again it can be a secondary diagnosis of dehydration.

And I've actually been looking at a number of medical records that I've been reviewing and looking at the nosologist which I have to say is one of my favorite words of all time. To see if they actually did.

Because these were folks who were receiving chemotherapy and virtually all of them did have that. Now, this is total anecdote and you don't make a measure based on anecdote. But they are going to be looking for those words because they'll get more payment.

So, when I first looked at this I was very concerned that they were going to be missing if it had to be the primary diagnosis. But with the secondary

diagnosis, all the complications for all the other medical problems, I felt like the coders are going to be looking at this because it is going to increase the reimbursement.

And it is a fairly easy thing to see. If the medical record from the ER says dehydration then they're going to put it in there even if it wasn't the final discharge diagnosis.

MR. VENKATESH: Actually, Patrick, can we ask you, is that true? That any patient with a secondary diagnosis of dehydration regardless of where it falls on the secondary lines is counted? Or is it only if they have one of those other primary principal diagnoses?

DR. SAMPSEL: Actually Patrick stepped away. And so Carol from AHRQ -- or Patrick, are you there?

MR. ROMANO: Yes, I am. I was about to step away.

Yes, I think that the latter part of your assessment is true. That you could posit a scenario under which appointment had a principal diagnosis of cancer and a secondary diagnosis of dehydration where they weren't counted. So that could happen.

I think that scenario is relatively unlikely because the coding rules are pretty clear that nobody wants hospitals to be billing for cancer care if what they're really doing is treating dehydration.

So, in most cases coding rules would require the dehydration to be principal and the cancer to be secondary, in which case it would be captured.

The exception would be, for example, a patient who was having chemotherapy and had chemotherapy-induced vomiting and then got dehydration from that. So in that case the dehydration would be attributed to the cancer and the treatment of the cancer. The cancer could still end up as the principal

diagnosis and it would be missed.

But that's a different situation from, for example, a patient who just happens to have cancer who experiences dehydration as a result of outpatient issues. Those patients generally would be captured.

So it depends on whether the dehydration is actually attributed to the cancer and the cancer treatment itself, or whether the cancer is an incidental diagnosis. Does that make sense?

So we would potentially be missing chemotherapy-induced vomiting is the bottom line if people are concerned about that.

DR. MCINERNEY: Marcel? No. Okay, are we ready to vote on reliability?

MS. ROBINSON-ECTOR: Vote is open. All votes are in and voting is now closed. Three voted high, fifteen voted moderate, zero voted low and one voted insufficient.

DR. MCINERNEY: Thank you. Okay,

validity.

MR. VENKATESH: So, for validity I think this is -- we've gone into some of this discussion during the importance discussion in terms of the primary concern being that I'm not sure if what we're measuring is what we think we want to be measuring in terms of the preventable dehydration visits within a hospital-based setting.

And the reason partly is one of the first things under validity it says is whether the specifications align with the evidence.

The evidence is really talking about things that can be done in the ambulatory setting to prevent emergency department visitation or the types of dehydration visits that we now call observation type visits. You know, shorter periods of dehydration. And so I think there's probably not as good of a link there.

They tested it at the score level and with a very similar model as before in terms of construct validity that looked I thought fine. The C statistic was just as high as it is for a lot of these claims-based measures. That seemed fine to me.

It's adjusted for age and gender. And the exclusions, just to note as a potential threat to validity especially in terms of what you think of in your head, it excludes all patients who are transferred from a skilled nursing facility. And so that may be a population a lot of people think of in their head in this dehydration group that could be preventable and have short hospitalizations. But they're actually out of this measure for similar reasons as to before that Patrick has already kind of highlighted.

What that means is I don't think it ends up throwing off the actual validity of the measure that much because two things

will happen. One is that it will make -- I mean I think that it just will be -- because they adjust for age while they were also excluding these events at the same time I think an areas estimate will be roughly about even. But remember that those are kind of out.

In terms of meaningful differences it's calculated very similar to the other measures. It shows that about 60 to 90 percent of counties currently do better than the 80th percentile with 5 to 16 percent that do worse than the 20th percentile. And so that would be potentially the counties with room for some improvement.

And then so ultimately I was left with kind of rating this as low. And the reason I rated it as low is that I think that despite all that testing I just am not convinced that measuring inpatient hospitalizations for dehydration is either linked to the evidence or that when you do

have that measure that it's meaningful to know what to do with it when we've missed the population that we wanted to measure in the first place.

DR. MCINERNEY: Further discussion on validity. Okay, let's vote.

MS. ROBINSON-ECTOR: The vote is open. Looks like one vote is missing. All votes are in and voting is now closed. For validity zero voted high, 8 voted moderate, 11 voted low and zero voted insufficient.

DR. MCINERNEY: Well, that brings us to a screeching halt on this one. The rules are we must have -- must pass both reliability and validity. We passed scientific acceptability and we did not pass validity.

MS. MUNTHALI: Yes, so what Adeela is saying, it's in that gray zone. So it's between 40 and 60 percent.

So I guess, you know, because of process we want to be consistent with what

we did yesterday with the other two measures that were also in the gray zone. I think the co-chairs can ask the committee whether or not they want to proceed and withhold voting. So I'll leave it up to you.

DR. SAMPSEL: Well, we can proceed through feasibility. We'll proceed through feasibility and usability as we did and then we could table the final vote.

DR. MCINERNEY: Okay. Arjun, would you want to take us through feasibility and usability, please?

MR. VENKATESH: Yes. I mean, for feasibility it's the same as the previous measures. It all uses administrative claims data, easily available and the software is available.

DR. MCINERNEY: All right. Any discussions on feasibility? Okay, let's vote on feasibility.

MS. ROBINSON-ECTOR: Voting is open. All votes are in and voting is closed. For

feasibility 15 voted high, 4 voted moderate, zero voted low and zero voted insufficient.

DR. MCINERNEY: Thank you.

Usability.

MR. VENKATESH: Usability. Again, the same thing. It's used in a variety of public health uses, national public reporting uses.

The only thing I noted was that one of the things listed on the form was CMS's QRUR reports which are Quality Resource Use -- I don't know what the other R is. But they're given to individual providers -- what is it? Resource Use and Quality -- oh, Quality and Resource Use Reports.

Anyway, the point is that this measure is again given to people around downstream payment policy and things like that.

And so one of my fears was that since one of this includes unintended consequences was that continued endorsement

of the measure in some ways signals that, hey, we think that there is some validity to this measure and I'm not so -- you know, I don't really love the idea of saying, hey, we should tell everybody what their dehydration admission rate is when we don't really know if that's what's being measured at all.

And then in terms of -- I think that was it. Yes.

DR. MCINERNEY: Any further discussion on usability? Okay, let's vote, please.

MS. ROBINSON-ECTOR: The vote is open. All votes are in and voting is closed. For usability 3 voted high, 10 voted moderate, 6 voted low and zero voted insufficient information.

DR. MCINERNEY: Here's the big one, overall suitability for endorsement. We can either delay that vote or we can vote now. What's the pleasure of the committee?

Delay? All in favor?

MR. STOTO: What will we learn if we delay?

MS. KHAN: The purpose of delaying is just so that when we get to public and member comment you can see what the comments are and then decide how you want to vote.

DR. MCINERNEY: Okay, so in favor of delaying? Everybody? All right. We will delay that vote.

Okay, folks, we are now on the last measure. Congratulations. That's measure 0281.

MR. FRANCE: Measure 0281 Urinary Tract Infection (PQI 12). Ron and I will be presenting this one.

DR. MCINERNEY: Thank you.

MR. FRANCE: Or just go ahead and repeat everything Arjun said for the last one almost. They're very similar. ACSC measure rate per 100,000 hospitalizations, in this case for urinary tract infection

with similar requirements for numerator as in the last, 18 years and older. Excludes skilled nursing facility transfers, hospital transfers.

And has its central view that UTI represents inadequate or delayed treatment for outpatient urinary tract infection.

So the first question that we would look to is this question of the evidence base for this outcome measure. And here again looking at our algorithms for outcomes we look directly to the question of whether there's some -- the steering committee views some rationale between outpatient treatments or linkages of health system's failure to hospitalization. Rather than looking at it as a process outcome.

Let me just mention a few things around the review. On the evidence side again it's one guideline that they referenced and it was really an outpatient UTI treatment guideline suggesting there

aren't clear care pathways about how UTIs should be managed.

I think in -- and I might turn to Marcel, but for the elderly it's a little bit less clear about the pathways for diagnosing and treating UTIs.

There was also a study that was mentioned from 1998 that showed that for ambulatory care-sensitive condition hospitalizations 10 percent of them were caused by UTIs.

Looking at the data for UTIs in general you'll note that there's this variation across age groups. Our working group talked a bit about how the under 65 have a relatively low rate compared to the over 65 and that there is this skewedness to the data. And there is variation across regions, ages and genders.

So, I'll stop there and see if anybody wants to add anything regarding the evidence question.

DR. MCINERNEY: Any discussion, further discussion on evidence? Yes.

MS. ASOMUGHA: Just a quick question. So, based on the assessment that you guys did with the evidence. So you're suggesting that there's not clear evidence as to why we need this?

MR. FRANCE: So, I think it's this issue that Arjun brought up before that when the outcome -- when you have an outcome measure, then you -- we should be focusing on whether there are processes of care that might be managed.

Is there strong evidence that shows that high-quality outpatient care processes leads to these reductions in hospitalization for UTI. There was not a body of evidence presented with this measure along those lines.

DR. MCINERNEY: Yes.

MR. SALIVE: So, I guess, you know, I'm not sure I fully buy into that theory.

I mean, I'm very strongly a proponent of evidence but I think that if I read it right on these ambulatory-sensitive conditions it's an outcome measure and you need a rationale.

So the rationale again is that most all UTIs should be able to be managed on an outpatient basis. Maybe not all. Maybe some percentage. I don't know what the percentage is. No one probably around here knows.

So, I don't think I buy that that is a lack of evidence that is damning to this kind of measure. It still could be a good measure.

So, you know. I'm not sure that that -- it sounds like you're making an analogy rather than that we should not like it. And I don't see that as being very convincing to me.

MR. FRANCE: That was not my goal. My goal was to state what was presented in

the evidence review which was of a guideline.

As a geriatrician I appreciate your expertise in this area and agree that the ideal is of course treating outpatient urinary tract infections leads to reduced hospitalizations.

I just would state that that broader question of showing the stronger linkage isn't as clear. And it's not required for the steering committee for us to pass it on evidence.

MR. BIALEK: Marcel, I think that the issue maybe gets raised again when we talk about the performance gap. So is there a gap.

And within the performance gap there is the issue of can you make a change that will have an impact. Right? That's part of the performance gap discussion.

DR. MCINERNEY: Okay. Any further discussion on importance to measure

evidence? All right, let's vote on that please.

MS. ROBINSON-ECTOR: The vote is open. All votes are in and voting is closed. For evidence 13 voted yes and 7 voted no.

DR. MCINERNEY: Okay, so that's above our 60 percent by about 5 percent. So I guess we can proceed. How about performance gap, please.

MR. FRANCE: So just to point out again that there does seem to be this variation across age groups as well as regions. So there does seem to be a performance gap in hospitalization rates across these different settings.

Our group in discussing this wondered whether there may be a future state where this is a performance gap that's focused more on the 65 and older cohort rather than this younger group. See if anybody else wants to add anything to that.

That maybe there's a future state where this metric is reviewed by AHRQ and submitted in some future review that is looking more at a segmented 65 and older metric rather than an 18 and older metric.

MS. ASOMUGHA: I was just going to ask when can we make that sort of -- that commentary that perhaps this measure would be better if the target age group was more focused on the elderly population where it seems like there's clear evidence that this would be useful. If that's something we could do now or we could do then.

DR. MCINERNEY: I don't think we can change it now. It would have to be another submission.

Is there someone from the measure development group that would want to speak to this at all?

MS. STOCKS: I don't know if Patrick is still on. This is Carol Stocks from AHRQ. We could certainly take that into

consideration. I think we'd want to do some testing first to see the impact.

DR. SAMPSEL: I would say this is one of those we'll capture in the meeting notes. The information will be given back to the developer. And they are listening and taking notes anyway. And it's something they can consider for the future. But we need to consider this measure as it is.

DR. MCINERNEY: Ron.

MR. BIALEK: Back to Eric's comment initially. When we look at the performance gap it's a demonstration of quality problems and opportunity for improvement.

The data that were provided by the measure developers was that there's an opportunity to reduce hospital admissions potentially by 10 percent. And the developers say that 10 percent is by having access to ambulatory care. And there was one study that related to that.

So, a couple of issues. One is the

opportunity is present but the percentage is relatively low. The developer really didn't offer a lot on how to impact that. And the access issue if they meant use of ambulatory care is a little bit different than access too because the population may have access but not the ability to use. And so I just wanted to raise those as far as the performance gap goes.

DR. MCINERNEY: Thank you. Any further discussion on the performance gap? Okay, let's vote on that. I'm sorry. Marcel.

MR. SALIVE: So I think one other gap is that it's going up. So I do think there is -- I mean, there's a temporal trend going up which suggests -- it may be more than 10 percent if you buy into the 10 percent potential improvement. And that's still a considerable amount of morbidity, that 10 percent. It's a very high rate.

DR. MCINERNEY: Okay, let's vote on

performance gap, please.

MS. ROBINSON-ECTOR: The vote is open. All votes are in and voting is now closed. For performance gap 4 voted high, 14 voted moderate, 2 voted low and zero voted insufficient.

DR. MCINERNEY: Okay. The next to vote on is priority. Who would -- you want to continue the discussion on priority, please?

MR. FRANCE: Here I am drawing a blank. I'm trying to think what's the best way to say it.

I think there was the sense that UTI hospitalization in that ideal framework is prevented with the high-quality outpatient patient care delivery system may face some of the same threats as -- this might be more about construct validity to a degree. Sort of what Arjun was saying with the whole dehydration issue.

As was mentioned there's this

increasing rate. There wasn't a real clear understanding about why those rates were going high. They said maybe it's because of increased emergency department use as an outpatient setting.

But there wasn't much data that they shared with us. They talked maybe obesity and diabetes incidence increases might explain it. It could be due to higher rates of resistance to antimicrobial therapy as another source of this increasing rate.

So is UTI hospitalization a high, moderate, or low priority for the nation and health and well being? I don't know that we came up with a strong answer. Again, maybe I'll ask others to comment in. Ron?

DR. MCINERNEY: Further discussion on priority? Okay, let's vote, please.

MS. ROBINSON-ECTOR: The vote is open. All votes are in and the voting is now closed. For high priority 3 voted high, 12 voted moderate, 5 voted low and zero

voted insufficient.

MR. FRANCE: All right, so from a reliability perspective they did a report on a variety of statistics. The ratios seem to suggest that it is a reliable, moderately reliable metric as it's constructed.

From a validity perspective while there may be these broader questions of construct validity they did do the modeling that we talked about this morning that seemed to align UTI inpatient hospitalization with characteristics of communities that are aligned with the infrastructure of the healthcare system.

So reliability seemed moderate. Validity seemed high as a specific measure but from a -- and moderate from a construct validity perspective.

DR. MCINERNEY: Any further discussion on the reliability? Okay, let's vote, please.

MS. ROBINSON-ECTOR: Voting is open.

All votes are in and voting is now closed. For reliability 7 voted high, 12 voted moderate, 1 voted low and zero voted insufficient.

DR. MCINERNEY: Thank you. Feasibility. I'm sorry, validity. I'm ahead of myself. Validity, sorry.

I guess we're all ready to vote on validity. Please.

MS. ROBINSON-ECTOR: The vote is open. All votes are in and the voting is now closed. For validity 4 voted high, 14 voted moderate, 2 voted low and zero voted insufficient.

DR. MCINERNEY: Okay, now we can go to feasibility.

MR. FRANCE: Yes, so here the workgroup all agreed that it's like other metrics generated by electronic sources in claims. So we consider it feasible. It's feasible.

DR. MCINERNEY: Further discussion

on feasibility? All right. Vote, please.

MS. ROBINSON-ECTOR: Voting is open. One vote is missing. All votes are in and voting is now closed. For feasibility 16 voted high, 4 voted low -- moderate, sorry. Moderate.

DR. MCINERNEY: Usability.

MR. FRANCE: Again as a group we thought that this was a measure that was being used and that it is usable.

It may run into the same questions around whether it's a reflection of the classic UTI hospitalization issue, or if these are a more complex complicated group of patients in the hospital now than they were when this metric was originally created.

In general though it is being used across counties and states.

DR. MCINERNEY: Further discussion on usability? All right, let's vote, please.

MS. ROBINSON-ECTOR: Voting is open. All votes are in and voting is now closed. For usability 8 voted high, 11 voted moderate, 1 voted low and zero voted insufficient information.

DR. MCINERNEY: All right. Here we are at the finish line. We now have the overall suitability for endorsement for this measure. Any further discussion on this? Okay, let's vote, please.

MS. ROBINSON-ECTOR: Voting is now open. All votes are in and voting is now closed. For overall suitability 15 voted yes and 5 voted no. For measure 0281 Urinary Tract Infection Admissions Rate (PQI 12) the measure passes.

DR. MCINERNEY: Well, great. Thank you, everyone. We made it through all of our measures. And now the reward is a nice one and a half hour presentation about harmonization.

(Laughter)

DR. MCINERNEY: I know. We need to do comments first and then we can do our one and a half hour presentation on harmonization. So can the operator let us know if there are any comments?

OPERATOR: To make a comment please press * then the number 1. And there are no public comments at this time.

DR. SAMPSEL: So, I'd just like to thank Patrick and Carol from AHRQ for joining us and answering the questions that we had. And Patrick for joining us so early from the west coast.

MR. ROMANO: Thank you very much. I'm back. It's been a pleasure to be part of the committee discussion. If there are other comments or suggestions of course we'll be happy to share those with the AHRQ team.

MS. STOCKS: Okay, thank you.

DR. MCINERNEY: Very good. So, NQF team, what are we doing now?

MS. MUNTHALI: Actually, Adeela is going to give the committee a background on harmonization. And I'll talk a little bit about gaps. But it will not be an hour and 30 minutes. We'll try and get you out before 3.

DR. MCINERNEY: Thank you.

MS. KHAN: Okay, thank you, everyone for your hard work. Now you'll just have to listen to me speak for another 15 minutes.

So I'm here to talk about harmonization and harmonization within the NQF portfolio.

Just a little bit of background. The quality landscape contains a proliferation of measures and some that can be duplicative and overlapping. They address the same conceptual measure focus and the same target population.

So this creates a lot of confusion in the field for people to interpret these performance results. And it also can

increase the data collection burden for providers.

So our goal is to standardize and align specifications and definitions for related measures that can help alleviate some of these problems. And when there is a sufficient amount of overlap we like to select a best-in-class measure to be the most appropriate measure for us to use.

So we actually wanted to go over this because we do have several related and competing measures in this project.

Just talking quickly about definitions. When we're talking about the measure focus we're looking at the target process, the condition, the event and the outcome. And for the patient population we're looking at the regular patient population, who was being measured.

So here's a table here that can -- it's an algorithm basically for telling you whether or not the measure is related or

competing.

Some of the principles for selecting the best among competing measures is multiple competing measures can be --

MR. SPANGLER: Adeela? I'm sorry, can you get closer to the mike?

MS. KHAN: Oh, sure.

MR. SPANGLER: Thanks.

MS. KHAN: So multiple measures can be accepted with adequate justification. NQF prefers the endorsement of measures that include the broadest possible target population for whom the measure is appropriate and indicated by the evidence.

NQF prefers endorsement of measures that assess performance for the broadest possible applications. So, for as many possible individuals, entities, settings and level of analysis for which the measure is also appropriate.

Just to continue, if a single measure cannot accommodate the inclusion of

all relevant patient populations or entities a second measure can be considered for endorsement in which case we would ask that the measures be harmonized to the extent possible.

When the best-in-class measure is not clear it's appropriate to endorse more than one competing measure. At the time of initial NQF endorsement NQF should identify analyses needed to conduct vigorous evaluation of the use and usefulness of the measure.

This information should be provided by the developers to support a best-in-class determination at the time of our three-year maintenance.

Until the clinical data from EHRs are widely available for performance measurement and reporting, endorsement of competing measures based on different data types can be justified.

Two measures may be needed to

achieve the dual goals of, one, advocating widespread access to performance results, and two, mitigating to performance measures based on clinical data on EHRs.

Some of the principles for measure harmonization. Harmonization should not stifle innovation and it should be ideally addressed before the measures are submitted to NQF. And all of the measures that we have in this project actually were -- the developers were told beforehand if they had a competing or related measure and they all did speak to each other.

The reason they weren't evaluated in this project was because they're out of phase. Some of the measures that are in the health and well being portfolio are not being evaluated during this phase of the project. So when they do come up again together that's when we'll tackle the harmonization issues. Which is why we wanted to just give you a primer as to why

that's important anyway.

Harmonization should not result in inferior measures. Measures should be based on the best measure concept and the best way to measure those concepts.

And then conceptual harmonization, whether the measures are intended to address the same focus and target population should be determined before harmonization of technical measure specifications. So, your definitions, codes and algorithms.

Harmonization should eliminate unintended differences among the related measures. When there's a decision not to harmonize the measures the value of the different conceptualizations and technical specifications should outweigh the burden imposed.

And the availability of standardized definitions and specifications that can be used across measure is a desired goal, but they often cannot be established a priori.

So measure harmonization efforts will facilitate achieving standard definitions and specifications.

So here's a list of the related and competing measures within the health and well being portfolio. The diabetes measure, diabetes long-term complication admission rate is related to 0272 Diabetes Short-term Complication Rate and 0638 Uncontrolled Diabetes Admission Rate.

2511 Utilization of Dental Services is also related to 1334 Children who Receive Preventive Dental Care which was developed by HRSA, and 1388 Annual Dental Visit developed by the NCQA.

We also have one competing measure in this project, 2528 Prevention Topical Fluoride for Children at Elevated Caries Risk. And that we found was competing with 1419 Primary Caries Prevention Intervention as Part of Well or Ill Childcare as Offered by Primary Medical Providers.

MR. KROL: I think I'm struggling with what determines competing. So, for that measure if I remember correctly the provider being measured in one is the dentist and on the other it's non-dentists. So, is it just the competition that they're providing a potentially similar service or partially similar service? I'm sorry.

MS. MUNTHALI: Yes, I don't know if we have the comparison table up. I think that would be helpful. Because what we do is spec it out against the data elements, the setting of care. So it's a side-by-side comparison. And I'm sorry, I don't have it in front of me.

But for one reason or another we made the determination that it did fit within the competing measures category because the -- it had the same concept for our measure focus and target process and in the same target population.

So I would be better able to -- I

don't know if you have the comparison. If you can bring it up we could show you where.

MR. KROL: That's okay. Just, I know they made an important point, at least I thought it was an important point of delineating dental services from oral health services. And the competition would lie in the oral health services side of things rather than the dental services side with those two.

MS. NISHIMI: Right, and then the -- the exception would then be because there might be a need for both of them that's why the committee would choose to apply that. But on its face for the other details they would be considered.

MS. MUNTHALI: And some of their explanation came about as a part of this process and not in their submission form.

So when we're initially determining whether or not measures are competing or relating, we are looking at their submission

form and then we get additional information, clarity from developers on what exactly they mean.

MR. BIALEK: And so the harmonization issue does not go beyond NQF? So Health People 2020 is off the table. Is that true?

MS. MUNTHALI: Yes. So our harmonization protocol and process is based on our endorsed measures. So a measure must be endorsed before we consider it -- NQF-endorsed before we consider it for harmonization.

But we could recommend to I think it was AHRQ when we were talking about Health People 2020. And we could recommend that they make, you know, they do align their measure with the denominator. And I can't remember the specifics of the conversation, but yes.

MS. BURSTIN: But just to add to that, I think that's a really good point,

Ron. And I think if you know from where you sit that there are in fact other standards out there that may not be part of our process, but others will be held to for other reasons that would be really informative I think to just bring it to us for information's sake.

Because the last thing you want to do is have people be forced to -- the whole point of this is to reduce burden and make sure people are being measured on what the quality signal really is and not just undue burden.

MR. BIALEK: And there are a couple of in the oral health area around sealants and use of dental services.

MR. STOTO: Are we being asked to do something at the moment about this?

(Laughter)

MS. KHAN: No, this is just a primer for the committee. Just because we never really get a chance to really delve into

what we mean when we say measure harmonization.

And so as part of a new organization-wide change we want to really up front bring the conversation about harmonization to the committee just so that when the time comes you're ready for the discussion.

MS. MUNTHALI: And it is a change. Sorry, Mike you were on previous projects. So what we did on previous projects was to start talking about harmonization and an action was required at the in-person meeting.

But we realized developers didn't have an opportunity to prepare their discussion points around harmonization, whether or not it's justified or not.

And then also as Adeela said some of the measures were not due for their maintenance review. So we were asking of them additional information before time.

And we realized that what we'd like to do now is kind of point out where we see some concerns over harmonization or where there may be competing or related measures, and then give them time to work that out. They've had those initial discussions as Adeela has said. And so we'll see where they go from here.

MR. STOTO: So, the ones that are already endorsed, they're endorsed. So the question is about the new ones, whether they can be more harmonized with that.

MS. MUNTHALI: All of them are endorsed.

MR. STOTO: I mean are you -- will you ask the people who have had a measure endorsed last year or the year before to reconsider it?

MS. BURSTIN: Yes. And that's the point of this is we will bring the portfolio to you. We'll have an opportunity. And they have then a year till their annual

update to harmonize, make sense of this.

We know they can't get it done immediately. Their processes don't work that way. We want to give them your best thinking.

And so a measure that's out there that's been endorsed for a year is just as much up for harmonization as one that's newly brought forward to you.

We just want to get the best-in-class out there, reduce duplicate measures that are slightly off. There's nothing worse than that as a clinician I think for many and others.

MS. MUNTHALI: And the measures in this project, of course we wouldn't be able -- it wouldn't be confirmed until they were endorsed. And you just endorsed them and so this is why we're bringing them to you. You've just recommended them for endorsement I should say.

MS. KHAN: Are there any other

questions? Eric?

MR. FRANCE: I'm sorry, just to remind myself, the competing. Weren't those two different bodies of care delivery systems, the first being dentists and the second being pediatricians and family physicians? And so are they competing in that regard or are they related?

MS. KHAN: They were classified as competing because they have the same measure focus and the same target population. The treating physician wasn't really taken into consideration.

MR. FRANCE: Thank you.

MS. KHAN: Okay, I can turn it over to Elisa to talk about gaps.

MS. MUNTHALI: Thank you, Adeela. This will be a very quick discussion just because of time. But also we would need quite a bit of time to talk about this.

But before we did start talking about the gap areas in our previous project

we talked about some of the concerns around trying to get the measures that the committee would like to see and NQF would like to see come through our process.

And for us at NQF we think that this is probably more important to tackle. We've been all through the MAP process, the measures applications process and our most recent project on the Community Action Guide and also this project. We've been talking about the gap areas.

And we did put a chart after this that kind of maps out the different areas that each group has mentioned. But we're really struggling about how we can get those sort of measures here, how we can connect with communities that may be using measures and see if they can come through the process.

These were some of the strategic recommendations that the prior committee came up with some of which we're already

tackling because they're global concerns that we have across multiple projects, not just the population health, health and well being project.

But I just wanted to share these with you. And we will probably follow up online or through our conference call to talk about gaps further. But I'll just go down the list.

One of the concerns or one of the ways that our committee, our previous committee said we could probably improve future calls for population health and health and well being measures is to identify population health measures that potential partners may be using.

So, to conduct a collaborative analysis of those partners which is essentially environmental scan.

We think that through the work that we're doing now through the Community Action Guide it's very different for NQF to be in

this space because we will be going out to communities, working with communities in addition to public health sectors and the clinical care delivery system to see what we can do to help improve population health. So that might be one way that we're doing it. But we'd welcome your ideas and suggestions on how we can move forward in this respect.

Another was really trying to refine guidance and definitions that we provide developers and our committees. Many have criticized us because they have said that the definitions and guidance that we've used have been very clinically focused. And so we may be talking the same talk as people in the communities but no one knows that because it's very medical-focused. And so they suggested that we come up with examples and guidance of what good looks like.

We are really happy to announce that since this project in 2011 we've done that

for all developers. It's part of the developer guidebook. You received a steering committee guidebook but we developed also a developer guidebook that includes examples of what good looks like. And so we'd like to move further here as well to see perhaps there are other examples for folks that are working on community health measures or population health measures.

Another criticism is that our submission process can be very odious and very burdensome. And so they had suggested to our committee that we provide more detailed technical assistance.

And this is something that we've done throughout all of our projects. We've added technical assistance as actually a budgetary line item in all of our project budgets.

And we worked very closely, for example, with the Dental Quality Alliance.

We worked with them for several months for them to prepare their submissions for this project.

They also suggested establishing an interactive community forum where there can be some sort of bidirectional learning and sharing. And this is something we're continuing to work through, talk about how we can do this.

But one of the things we do have, we've had for many years is monthly developer webinars. And so population health is one of the topic areas that we're talking about, not just that we have this work at NQF but also talking about some of the methodological challenges to population health measure development.

And many thought that NQF really needs to confirm what their value proposition is to folks who are working in communities, people working on population health. We are working towards that.

As Helen has mentioned it is a very large strategic goal of ours to improve and strengthen population health work. And we're doing it through the different projects that we've all mentioned over the last couple of days.

And so I don't know if you have any other suggestions of what your general feedback is on some of these strategic goals that were set about two years ago. Many on this committee were on that committee. And so I don't know if you would like to share your perspectives.

MR. STOTO: I was and I think these were important. I especially think the last one is important. Because I think that as much as people have come to realize the value of endorsement by NQF in the healthcare sector people in the public health sector have no idea even what it is. And so.

And I think that one way that we can

address that is in the area of community health needs assessments, that hospitals are now required -- that non-profit hospitals are all required to do. And the health departments are required to be part of that, although that doesn't always happen as well as we'd like.

And you know, if in the IRS guidance that says what is an acceptable community health needs assessment is something about NQF endorsement being important that would make a tremendous amount of difference. Right now the IRS guidance says almost nothing about what makes an appropriate measure. It says nothing about what makes an appropriate measure for a community health needs assessment.

DR. MCINERNEY: The United States clearly lags behind most of our European countries as far as public health and healthcare quality and quality of health of our citizens compared to countries in

Europe.

And I just wondered are there some things that are going on across the pond as they say that we may be able to learn from what's happening in those countries that can help us inform how to improve our quality, one of the three measures there, the healthcare quality of our population as a whole? I wonder if folks have thought about that.

MS. BURSTIN: Yes, so we've actually got some projects where we've actually been trying to get some of the learnings from the other countries across the pond.

And in fact have a project right now we're doing on gaps on patient-reported outcomes which has been directly, for example, looking at the PROs already in use in the UK.

So again, as we think about this going forward examples of international measures that we should consider bringing

forward in the U.S. would be very welcome as well.

I mean, we often hear about how some countries like Sweden have 10 measures they use to manage their health. That would be glorious in comparison to the hundreds that we have in Health People 2020.

And then also just another thing to keep an eye on is the fact that the IOM has a committee right now on coming up with a core set of measures.

And at least from those I know on the committee, not who can share significantly, imply it's again at a fairly high level. I think it will be more at the population level as well.

So, I think there's a lot of things happening -- that's coming out in August -- that we should keep an eye on.

MR. BIALEK: Just a couple of thoughts. This is highly conceptual for many of us in public health. And if there

are any examples of users who are not clinical providers that would be enormously helpful as well as examples of the actual measures of what's released from NQF. So that would enable us to take that to other organizations who may be willing to develop measures.

A second item has to do with the time frame which tends to be fairly short. And even, you know, we heard from our colleagues from AHRQ who did a magnificent job how difficult it was to put all of that together in the time frame.

And I think when we're reaching out to non-governmental organizations or organizations that are not quite as large as we may be accustomed to, the time frame maybe needs to be extended.

MS. BURSTIN: And part of our goal as well in our discussions with CMS about our contract going forward will be not to have sort of this rush of deadlines for

these various projects, but instead to have a more smoothed out process. So we're piloting that currently in our endocrine project where measures can be submitted every six months to the standing committee.

Whether that will be every six months or annually it would allow people to say it's not this urgency, oh my God, if I don't get it in now I can't get it in for two or three years. So I think that's our vision for that, Ron, and I think that would help.

MS. NISHIMI: And, Ron, I'm going to correct the record here because I'm a consultant and I can do it on behalf of NQF.

But AHRQ was actually given a pass two years ago and they've known for a year and a half that they had to meet these deadlines. So, the comment by the developer that they had 48 hours was entirely the developer's doing.

MR. VENKATESH: I guess I have two

sets of comments.

One is conceptually I think we sometimes miss the title of the committee which is Health and Well Being. And the vast majority of measures we evaluate -- and that's reflected in the vast majority of the measures we have are really of sickness and not well being, right? And that is largely a construct of data that we have available too.

But it would be nice if in the future the way the strategy is set up is in a way where it thinks about what actual well being measures, like what type of functional status measures might there be for high-risk populations and things like that where you're actually health-promoting measures as opposed to the avoidance of bad events.

And along those lines I think that the challenges that the data right now that we've traditionally used is a lot of administrative claims or even things that

we're going to pull from electronic medical records are still going to capture bad things.

And so in that vein at least I came up with three ideas of things I think would be good community-level measures and address -- one thing we hadn't done is we didn't really look at what we have here across the National Quality Strategy by domain.

And two domains that stuck out to me as missing where the community level actually may be hopefully a less politically charged place to measure and may actually be a better place to measure around care coordination and affordability.

And so for care coordination, a couple of measures that came to mind is why are we not measuring community-level risk standardized readmission rates.

The argument from the hospitals has been this is largely driven by a lot of community resources and factors. And we

have a lot of other things, some research that says communities may drive that.

I'm biased because we make these measures. But I think that's one type of measure.

Another would be we haven't really gotten into the behavioral health space a lot with a lot of measures within this group. And so I was thinking about something that would measure days in emergency departments or observation stays prior to a psychiatric hospitalization as a measure of mental health care access in a community.

And then the last one I thought about was something that we're trying to deal with in New Haven is that on the affordability side for communities we really think about that in the municipal level. And healthcare costs are bankrupting our community. And that's really spending on public employees' healthcare benefits.

And so we have measures of total spending per beneficiary. What if we applied that to all public employees within a community? And thought about how much we were spending of community dollars within the community on healthcare.

And so I think there's ways to start going -- those are all kind of controversial ones. But I think that we should be thinking about each of those domains as well as also health measures.

MR. CARILLO: Yes, in the past two days perhaps a leading recurring theme is the social determinants of health. And I think that they apply not just to the population health space but they apply across the board and practically with any health measure you can think of. So I wonder to what extent that's something that NQF can focus on.

I think, for example, our PQI measures that we went over today, if we had

parallel measures of social determinants you might be able to get a better sense of what the measure is telling us.

I mean, we were uncomfortable with several of the PQI measures because of that, because of all the confounding. And frankly, I mean the committee was almost uniformly and unanimously concerned on those bases. So I think that that's an important priority that I would recommend.

MR. STOTO: Coming back to the question of international comparisons. A number of European countries have something called population health observatories. And I don't know much about them but there's a network of them. And it might be interesting to see what kind of measures that they are using in their work.

MR. SPANGLER: I wanted to go back to Arjun's point. Because this came up when we were on the previous committee about population health. Can we get health

measures.

And one of the things that we kept hearing from the developers is we can't do it. They're too difficult to develop. They cost too much money. They're too expensive. It can't meet the criteria that NQF has.

And I know you guys have tried to change the criteria a little bit to adapt to that, but I still think that's an issue with measure developers.

And also, historically most measure developers, they don't know how to develop these type of measures. So we need other people to kind of come in and help them do that.

And I know that that's a process that's going to take awhile. But I think some of those issues that we had discussed previously are still there.

MR. STOTO: You know, the Behavioral Risk Factor Surveillance System has data on healthy days that's used quite widely. And,

you know, we used other BRFSS data all the time in this kind of stuff.

And you know, if our standards are such that that doesn't pass then maybe we should think about changing the standards.

That's out there, it's being used all the time. The county health rankings do it for every county in the country every year.

MS. BURSTIN: And just one more comment to add into that. One of the things we've also been talking a lot about is whether we've moved beyond the days of a binary yes/no for NQF endorsement, and whether endorsement should be more related to fit for purpose.

So if a measure is intended, for example, for population surveillance, or QI versus payment, would you have different requirements even among testing. I mean, what kind of testing do you need if you're measuring a community who may not have a

pay-for-performance attached to it or public reporting?

It gets complicated. In the past when NQF has done this years ago, before I came, in one particular project around cancer. The concern we heard from a lot of the consumers and purchasers was that some of the measures they cared about most wound up in the QI-only surveillance realm. So we have to be really careful of what that looks like.

But I think it really goes all the way back to measure development. Sometimes there are measures that just will be developed differently if your intended purpose is not for some of the sort of high-stakes financial accountability where misclassification has a pretty significant effect.

But also public reporting. I mean, the last thing you also want to do is have somebody go to a website and find

information that's incorrect because the measure is not sufficiently valid.

MR. SALIVE: I also agree about measuring health and think that we should try to do it. And I agree that functional status is a good idea there, quality of life.

And on the flip side I think I have some thoughts about the older population and multi-morbidity is another thing which I think has been ignored both by the health system and by population health.

And we touched on this today I think with the exclusion of nursing home patients from a lot of these measures, saying that they're measures of the community. And I think that it disenfranchises the nursing home people really by doing that, that they have these issues. And they're in the healthcare system, kind of, and they're definitely living in the certainly.

It's a fine public health transition

that they've excluded them from the denominator. And so then you argue well, they can't be in the numerator either.

But I think that begs the question of why and we need to fix that I think sometime down the line.

I also had one final technical comment which is just that bullet that said something about the value proposition to me is very cryptic.

And so I think that whenever you talk about value, I know it's a loaded term, so value to who. And who is spending the money on that.

Is it really just that your measures will be used, is that what this is about? Because I would be in favor of that.

Or is it about that we should have measures that deal with economics?

So I guess I would just -- I think those are good ideas, that we should flesh this out a bit and not leave it in this sort

of crypto-speak. You know, we don't want to say we're going to save money or have death panels, I'm sure.

MS. MUNTHALI: No, I think this was really focused on engagement with people that have typically not engaged with NQF. So, people outside of the clinical care delivery system. And really, what is the value of them coming through our endorsement process.

And I will actually share with you the final report. I think the committee did a fabulous job and their recommendations were spot on. I think it will provide additional context.

MS. FRAZIER: As I said earlier or yesterday, one of the reasons, probably the only reason that I did this committee is because I really want to help with the transition out of the medical mode in how we look at health and well being.

And I think that a couple of things

I've heard that make me feel that we can move that transition. It's going to take time.

One is an openness of NQF to re-look at their measurement process. To be open to new ways of looking at measures and how we effectively do that.

I think two is the research, all the best practices we can all over the country. Because I think there are people doing things that we're not even thinking about when it comes to health and wellness.

I'm reading material that is fascinating and innovative but it doesn't fall into the box that we're used to. So I have to open up my mind as well of how to do this.

I think third is to take advantage of tools that we already have. When I look at the behavior risk factor I don't love it, but I think was it applied more in local communities it may be more effective.

The way it's used now is so broadly -- the way the CDC uses it is just not as effective. It's difficult to utilize it in a way to be actionable. So I think it is a tool, I agree with you, but I think it has to be reapplied differently.

So I mean those are the three thoughts I have. But I think this is going to take some transition.

And I think one of the things that this committee can do is maybe getting some things to read to try to begin opening up our minds to figure out how we do this differently. So that would be my thoughts.

DR. SAMPSEL: And I would just follow up on that. And I can help try to pull some things together.

But I was recently working on a project with a hospice expert. And it was talking about measure development in hospice.

And she had some comments about

measuring patient priorities. And I said well how do you do that, what's the evidence? She said well quit thinking like a measure developer and you'll figure it out.

And you know, I haven't been a measure developer for five years now but it's still hard for me to think outside of the evidence, you know, testing and all of those features.

So I agree, it's stepping outside of the box, learning best practices, because there are some opportunities out there. And I think one such thing is PROMIS. Just really kind of an opportunity to think differently.

And the other area I think of great need, it's hard to figure out where it fits though, is that quality of life that I think I've heard a couple of folks mention, that everybody is struggling with.

How do you measure quality of life

and what does quality of life mean based on where you are on an epidemiological scale?

MS. MCKANE: I just wanted to add that I've been kind of the survivor of a number of different indicator projects that are ongoing as that's a very popular thing. We're epidemiologists and we're in public health so we have to measure everything.

And I think that there is a place for NQF and population medicine to work together on this. Because when I see these indicators, the quality varies, the criteria varies. There's really, there's a lack of consistency. And it would be really nice to have a consistent source that you could refer to.

And actually we often are, you know, we go through the NQF indicators. We go through the other ones that are online that have been through a vetting process and we're trying to look at things. So I think that there's certainly -- I think there's a

value in that. And I would like to see it proceed. I do think some of the hurdles that everybody has described we need to work on.

And the other thing I have to point out because I am an epi and we do work with BRFSS data is that we don't really particularly care for the county health rankings and the methodology. It's not really been approved by CDC. We kind of -- we take the data and do the modeling.

But -- and CDC is working on developing methodology to improve -- to develop a consistent method of developing county-level data. We actually do it in Michigan at county level.

It depends on the population because smaller areas, smaller counties may need more years of data. And it does affect the validity.

MR. FRANCE: Last word. Just to remind us that health occurs more in the

communities than in our clinics, that it's where we live, work, learn, play, pray. And so those are all domains that we might want to think about, work site, school site, family health.

And not be afraid of the fact that that feels kind of weird to think that way. But that's where health and well being lives.

MS. MUNTHALI: Thank you so much. This has been very valuable input. We are so excited about the standing committees. I know your sentences are two to three years.

(Laughter)

MS. MUNTHALI: It gives us an opportunity to really engage with you and to share ideas and move forward on population health.

And we just wanted to really thank Sarah and Tom for their leadership. You've done a great job. The entire committee has done a great job.

We'd also like to thank the developers, NCQA. We'd like to thank AHRQ and also the Dental Quality Alliance. And everyone else who's listening, thank you for participating.

DR. MCINERNEY: Thank you to the committee. You really were a great group, very cooperative, collegial and a lot of good input. And thanks to the staff for all of their help in making this be so successful.

(Whereupon, the foregoing matter went off the record at 3:01 p.m.)

C

P

T

C E R T I F I C A T E

This is to certify that the foregoing transcript

In the matter of: Health and Well Being
Steering Committee Meeting

Before: NQF

Date: 04-30-14

Place: Washington, DC

was duly recorded and accurately transcribed under my direction; further, that said transcript is a true and accurate record of the proceedings.

Neal R. Gross

Court Reporter