NATIONAL QUALITY FORUM

+ + + + +

POPULATION HEALTH ENDORSEMENT MAINTENANCE
PHASE II
STEERING COMMITTEE MEETING

+ + + + +

WEDNESDAY
MAY 30, 2012

+ + + + +

The Steering Committee met at the National Quality Forum, 9th Floor Conference Room, 1030 15th Street, N.W., Washington, D.C., at 8:30 a.m., Paul Jarris and Kurt Stange, Co-Chairs, presiding.

PRESENT:

PAUL JARRIS, MD, MPH, Co-Chair
KURT STANGE, MD, MPH, Co-Chair
RON BIALEK

SARAH LINDE-FEUCHT*
LINDA KINSINGER*
MADELINE NAEGLE
ROBERT PESTRONK
SUE PICKENS
MARY PITTMAN*
AMIR QASEEM

SARAH SAMPSEL
JASON SPANGLER
MATT STIEFEL
MICHAEL STOTO

NQF STAFF:
HELEN BURSTIN, MD, MPH
KRISTIN CHANDLER, Project Analyst
ANGELA FRANKLIN
ELISA MUNTHALI, MPH, Senior Project Manager for Population Health
ROBYN NISHIMI, NQF Consultant

ALSO PRESENT:
DAWN ALAYON, NCQA
SAM AMIFAR, New York City Department of Health and Mental Hygiene*
MARY BARTON, NCQA*
SEPHEEN BYRON, NCQA
IAN CORBRIDGE, HRSA*
ERIN GIOVANNETTI, NCQA
IRENE HALL, CDC*
SHARON HIBAY, Quality Insights of Pennsylvania
PEGGY HONORE, Office of the Assistant Secretary for Health, HHS
DAWN JACOBSON, Public Health Institute*
BOB REHM, NCQA
ALFONSO RODRIGUEZ, CDC*
AMISHI SHAH, Office of the Assistant Secretary for Health, HHS

DON WILSON, Quality Insights of Pennsylvania
PASCALE WORTLEY, CDC*

*Participating via teleconference

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

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

9:02 a.m.

CO-CHAIR JARRIS: Good morning everyone. This is Paul Jarris. I'm here with Kurt Stange and Helen Burstin and Elisa Munthali for the day. And we have a number of people in the room we'll introduce in a moment.

Thanks for the traveling. And for those of you who were unable to travel, thank you for joining us by phone. So we do have about a day and a half or so of meeting to go over.

Like in the past, the first meeting will begin with looking at some of the measures that we received. And very importantly, I think, this is an opportunity to look at what we received as compared to perhaps our expectation, as well as compared to the NQF population health measures criteria.

And then tomorrow we'll have more

discussion to talk about. For the status check where we are, the progress we've made, and compare that to our expectations to what we received.

And really have a more in depth discussion, both with measures developers as well as the committee on the steps going forward. How should we continue? What types of things should we do to really increase the participation in this process, and in the yield on the measures?

So I think it should be a very fruitful discussion. And it could be very important tomorrow to provide guidance, both to the committee, the workgroup, as well as NQF in moving forward on population health measures. Kurt? Okay. Well go ahead, Elisa.

MS. MUNTHALI: Good morning everyone. This is Elisa Munthali. I'm the senior project manager for population health. And I just wanted to welcome and thank you again for joining us today.

I just have a couple of housekeeping items before we continue with our introductions. As you can see, our meeting is being taped and transcribed. So we do ask that whenever making a comment or posing a question, that you make sure you turn on your microphones and speak into the mics so we can capture your comments.

And as for the restrooms, a very important item for all of you that are here. They're just in the hallway, just beyond the elevators to the right.

And so what we're going to do are introductions of everyone who's in the room. And for everyone we ask that you state your name and your organization. And for the steering committee, we also ask that you give us a little bit more about your role within the organization.

As you recall during our first in person meeting, you did publicly announce your disclosures. We're also asking now if there

are any updates to that disclosures, that you do that as well. And that includes any changes in employment. I think that is it with regards to the disclosures.

I wanted to bring up also the measures that are under consideration. And part of the disclosure update is to consider these measures and to see if you've had any involvement in measure development. And so as you go around, if you can please indicate so if you've had that. So I think we'll start to my left with Paul.

CO-CHAIR JARRIS: I'm Paul Jarris, Executive Director of the Association of State and Territorial Health Officials. And I have no conflicts on either current or past to announce. Thanks.

CO-CHAIR STANGE: Kurt Stange. I'm a professor at Case Western Reserve University in Cleveland, and also working for the year with the implementation science team at the National Cancer Institute. I have no

new disclosures. I have been involved in some measure development in the remote past. But nothing related to this topic.

DR. BURSTIN: Good morning. Helen Burstin, NQF.

MEMBER STOTO: Good morning. I'm Mike Soto from Georgetown University. I don't have any conflicts to mention.

MEMBER NAEGLE: Good morning. My name is Madeline Naegle. I'm a professor at New York University. There my primary focus is on substance related disorders programming development.

I serve as a Deputy Director on our Center for HIV and Drug Use Research. I'm also a co-investigator of Project SARET, which is substance abuse research training and education for our students. Good morning. I've no disclosures.

MEMBER BIALEK: Good morning. Ron Bialek, President of the Public Health Foundation. And I have nothing new to

disclose. As far as measure development, I had some discussions with the Legacy Foundation regarding the tobacco measures.

MEMBER SAMPSEL: Good morning. Sarah Sampsel with WellPoint. My position has changed at WellPoint. I'm now Business Change Director for our Clinical Quality Division, which is a group that's actually trying to improve performance measures across WellPoint. And I should disclose that I was lead staff at NCQA for the development of measure 0024 and 1690.

MEMBER PESTRONK: I'm Bobby Pestronk, the Executive Director of National Association of County and City Health Officials. My position hasn't changed since the last meeting or disclosure. I don't have any conflicts to report. And I have not been a contributor to the development of any measures under consideration.

MEMBER SPANGLER: I'm Jason Spangler. I'm the Chief Medical Officer of

Partnership for Prevention. I don't have any disclosures. Are we supposed to just talk about the measures there that we've possible worked on?

MS. MUNTHALI: Any of the measures under consideration. So also in Workgroup B.

MEMBER SPANGLER: Yes. I haven't worked on any of those. But kind of like Ron, I don't lead our tobacco efforts. But I work with the people on some tobacco measures.

MEMBER STIEFEL: Matt Stiefel from Kaiser Permanente. Senior Director in the Care Management Group for population health. I haven't worked on any of these measures in previous groups. And no disclosures.

MEMBER QASEEM: Amir Qaseem, American College of Physicians. No financial conflicts of interest. But non financial, I'm on various boards and committees, just like everyone. And a bit of work with CDC and all of that. I don't think that's something.

MS. CHANDLER: Kristin Chandler,

NQF project analyst.

MS. NISHIMI: Robyn Nishimi, NQF consultant.

MS. ALAYON: Good morning. I'm Dawn Alayon. I'm from the National Committee for Quality Assurance.

MS. HONORE: Peggy Honor, Office of the Assistant Secretary for Health, HHS.

MS. SHAH: Amishi Shah from Office of the Assistant Secretary for Health at HHS.

MS. FRANKLIN: Angela Franklin, staff NQF.

MS. MUNTHALI: Sarah, on the phone, and Linda.

MEMBER LINDE-FEUCHT: Yes. Hi, this is Sara Linde-Feucht. I'm the Chief Public Health Officer at the Health Resources and Services Administration with HHS. And I have no disclosures, and have not worked as a developer of any measures.

MEMBER KINSINGER: And good morning. this is Linda Kinsinger. I'm the

Chief for Preventive Medicine for the Veterans Health Administration. And have no disclosures to report, and have not worked on any of the measures.

MS. MUNTHALI: Thank you. Is there any other steering committee members on the phone? Okay. Thank you everyone. And so before we turn the meeting over to the committee and to the co-chairs, we do have a few introductory slides, some reminders.

We just wanted to first tell you why we're meeting today. And during this meeting the steering committee will evaluate endorsed and newly submitted clinical body mass index measures and population level measures against NQF's criteria and guidance for population health, to determine their suitability for endorsement recommendations.

They will also identify related and/or competing measures to determine if there's need for harmonization or selection for best in class.

And finally, they'll participate in discussions about outreach efforts for our recent call, provide feedback on improving response to future calls for measures, and ways in which to strengthen NQF's population health portfolio.

The committee will review and vote on nine submitted measures that are under endorsement consideration. Developers are present, as we noted earlier. And they will briefly introduce the measures at the start of each topic area. And they're also available to respond to questions from the committee.

The committee in turn will discuss and vote on each of the measures using the four major criteria, importance to measure and report, scientific acceptability of the measure properties, usability and feasibility.

Our members and the public will have several opportunities during the day to provide comment. During Day 2 the committee will participate in several discussions,

including a facilitated conversation with invited developers on how NQF could improve future response to call for population health measures.

Later on in Day 2 the committee will look at NQF endorsed population level measures, and offer input on how to address the measurement gap.

And finally they will look at the recommendations from the commission paper to see if there are any they would like to adopt as committee's recommendations.

As with Day 1 our members and the public will have an opportunity to comment. We just wanted to remind everyone about the project goal and scope.

The population health project sought to expand NQF's current portfolio, preventive services and immunization members, which we did in Phase I. And healthy lifestyle behaviors and broader population health measures in Phase II.

And while we're doing this we wanted to also foster harmonization of these measures at all levels of analysis, including the provider and population level. And as a reminder, this project is being conducted in two phases. And this is just an outline of our endorsement process.

And as you can see, we're in this third stage of the process where the committee is reviewing and evaluating measures, and will render recommendations for endorsement. And so we wanted to go over the endorsement maintenance process with you.

And actually we have done this. So we're going to skip through these two slides, and just go through the measurement evaluation process with you.

Before turning over the meeting, we wanted to tell you what we did with the measure. We divided the measures into two workgroups. There's Workgroup A, which is the clinical body mass index measures. And

Workgroup B that's looking at the population health measures.

Several committee members have been assigned to each measure. And they will lead discussion today on those measures. However, we wanted to stress that all committee members will evaluate all measures.

The committee will also discuss and vote on each measure by criterion. And then if a must pass criterion, like importance to report, measure on report is not met, then the decision and discussion for that measure ends.

If the measure meets all of the criteria the committee will vote on the suitability for endorsement. We do have a cautionary note.

If they're related or competing measures, the vote is not final until there's review of whether or not harmonization is needed, or a selection of best in class is needed. And so I'll turn it over to Kristin

who is going to go over our electronic voting process.

CO-CHAIR JARRIS: Elisa, can I ask a question?

MS. MUNTHALI: Sure.

CO-CHAIR JARRIS: The clinical BMI seems to me to be back to the Phase I. Is there a reason it was considered here in Phase II, rather than in Phase I?

MS. MUNTHALI: Because it's suggesting healthy lifestyle changes. And that was part of the call for measures.

DR. BURSTIN: And I think actually our hope was that we were going to get measures in at the population level on obesity, that we would logically want to look at together with these.

So we actually had these already earlier on, but wanted to at least have them reviewed. We were surprised we didn't get anything in this area population wise.

CO-CHAIR JARRIS: Yes. That's why

-- Because the measures don't appear to be population wise, they are clinical. That's what I thought they were back in Phase I, which is fine. It's just they're really not population measures. But we'll talk about that when we get there I think.

MEMBER QASEEM: Sorry. You may already mention it to the committee. Where do the Phase I measures stand? I think the members have already voted now, right?

MS. MUNTHALI: They have. And actually they're in appeals right now. And so they have been endorsed. Nineteen measures were endorsed. I think 11 immunization measures and eight preventative services measures.

Those were the screening measures. And the appeals process will end on January 1. I mean, sorry, June 1. So we actually haven't, knock on wood, received any appeals so far.

MEMBER QASEEM: So then it will go

--

MS. MUNTHALI: But we'll let you know. Yes.

MEMBER QASEEM: And then it will go to CSAC, right?

MS. MUNTHALI: It's already gone to CSAC.

MEMBER QASEEM: Oh, it's already gone to --

MS. MUNTHALI: It's gone to the board.

MEMBER QASEEM: Oh, okay.

MS. MUNTHALI: Once it's gone through board ratification they're endorsed. Yes.

MEMBER KINSINGER: Elisa, for those of us on the phone, are the slides available somewhere?

MS. MUNTHALI: Yes. Kristin is going to send them to you. Is this Linda?

MEMBER KINSINGER: Yes.

MS. MUNTHALI: Okay.

MEMBER KINSINGER: Thanks very much.

MS. MUNTHALI: And she'll send them to you as well, Sarah.

MEMBER LINDE-FEUCHT: Many thanks.

MS. CHANDLER: So very briefly, most of you should remember the voting process from Phase I. But all of you should have received a voting remote when you walked in the door. If you didn't, please let me know.

I will begin the voting on each of the criteria for each measure. I will begin the timer and you'll have 20 seconds to indicate your response. You should point your remote towards this computer right here.

A green light indicates your vote was received. A red light indicates there's a problem with the remote. And in that case we'll need to exchange it out for a new one.

If you need to change your vote for any reason, just simply press the number that correlates with your answer. Only the

second, or last vote will count towards your, towards the total tallies.

So any questions? Yes. And for those on the phone, if you would like to verbally indicate your responses for each of the criterion as we go through the measures, we will record those here.

OPERATOR: To ask a question press star and the number one on your keypad.

CO-CHAIR STANGE: And Kristin, we're just voting on the four summary kind of measures, not the individual ones?

MS. CHANDLER: No. We will be voting on each of the major criteria.

MS. MUNTHALI: And Kristin will go through an example so we can get a hang of it.

MS. CHANDLER: So real quickly, if everyone could pick up their remotes. I'm going to start the timer now. And please enter any response. This is just a practice. Towards this computer. I think we're -- Okay. All righty? Okay.

MS. MUNTHALI: Great. So we will start in reverse order. We're going to start with the Workgroup B measures. The first on is Measure 2014. This is Place of Birth.

This is from the CDC. It's a new measure. The measure developer, Alfonso Rodriguez, is on the line. Operator, can you make sure that Alfonso's line is open? Anika?

OPERATOR: His line is open.

MS. MUNTHALI: Okay. Alfonso?

DR. RODRIGUEZ: Yes. Hello. Can you hear me?

MS. MUNTHALI: Hi. How are you?

DR. RODRIGUEZ: I'm very good, how are you?

MS. MUNTHALI: Fine, thank you. If you could go ahead and provide introduction? You can do the introduction for Measure 2018 as well. About three to five minutes?

DR. RODRIGUEZ: Yes. And thank you very much for the opportunity to propose

about these measures, this measure. I work for the Center for Disease Control and Prevention. And the first measure is place of birth.

And we are proposing this measure from three different point of view. One, we consider that country of birth is an important determinant of health. Because it's associated with barriers to access to care.

Are different in the various experience by the general population of the U.S. Also foreign born and migrant populations experience legal barriers to access to care, language barriers, and social isolation.

And a higher potential for discrimination and xenophobia than the general population or other minority populations in the U.S. And also they tend to suffer more, experience more health literacy challenges.

So we consider it's an important determinant of health. Also we're looking at

these variables from the vulnerable populations, sub-populations' point of view.

Foreign born populations, by definition, have experienced multi-national exposures. They come, many of them come from developing countries where the burden of disease is different than the one in the U.S.

And as a consequence they also show different health conditions, or a different prevalence of health conditions than the general population of the U.S. They also have different behaviors, attitudes and beliefs about diseases and access to care.

But also have different on the general, maybe different on the general population. The foreign born population is an increasing minority in the U.S. Thirteen percent of the U.S. population is foreign born.

Currently it's a highly diverse population. But also the children of the foreign born, who are born in the U.S., can be

also vulnerable population.

And the parents are the ones obviously who are making decisions about access to care and treatments, et cetera. And this is an increasing population, the children of the foreign born.

And finally, we believe that the country of birth is an important -- And this has been demonstrated in the literature. It's an important variable to measure to better assist and identify health disparities.

It will help to identify some groups within the race/ethnic groups that are more vulnerable for specific health conditions, including -- Or health issues, including health care, health insurance and quality of care.

There are many disparities in access to health services experienced by the foreign born, including infectious disease testing, like HIV testing, cancer screening.

Prevalence of the infectious

disease like tuberculosis and HIV are good example, hepatitis B. Chronic conditions, like certain types of cancer, like cervical cancer and others.

So we think it's a very important variable to help us to go deeper within the race/ethnic. So that would be to identify race/ethnic groups who have increased disparity for a specific health condition.

There are some -- This variable has been validated, has been collected by census in other, a number of national data sources for a number of years.

It's a highly valuable, reliable measure used to comply with international recommendations to increase in monitoring of the health of foreign born and migrant populations.

It's been collected, as I said, by a number of national surveillance and data sources in the U.S. But not by all. So one of the committee members mentioned about the

importance of standardizing, you know, the collection with national data sources.

And we believe that it's important to have these variables in every national information data source. So I don't know if there's any questions. I don't know how much time -- Hello?

CO-CHAIR JARRIS: Any questions?

OPERATOR: To ask a question, press star then the number one on your telephone key pad.

CO-CHAIR JARRIS: Anyone here in the room have questions? Amir, I think this was -- Or no, Alfonso's --

OPERATOR: At this time there are no questions.

CO-CHAIR JARRIS: Oh, Bobby? Was this --

MEMBER PESTRONK: So our job is to share with the rest of the committee the discussion that the subgroup has had. It is on, yes. Our job is to share with the rest of

the committee the discussion that the subgroup of us had about this particular, this submitted measure.

And based upon our understanding of the rules of review, it is essential for the measure to pass these, this first group of criteria in the impact opportunity and evidence section in order to move on and consider the other aspects of the measure.

And I think following the discussion and review, the group's consensus is that this is an important demographic characteristic, which is useful to know about certain populations.

But it doesn't rise to the level of a population measure specifically. It may be important data to collect as an aspect of some other measure. But it would be difficult to see this particular proposed measure on its own in the context of opportunities for improvement.

One wouldn't necessarily only want

to know whether place of birth, or the numbers of people, or the rates of people born in certain places had changed over time.

So understanding how the measure as proposed itself could help to, help with an understanding of a particular performance gap, I think there was not enough information, not enough explanation that accompanied the materials to have us pass that particular criterion.

And I think that because in those first two areas of impact and opportunity for improvement, the recommendation or the -- I don't know whether we're recommending to the rest of the committee, or just sharing with your our review and analysis. Is that this doesn't meet the criterion to have, to accept the measure as a population measure.

DR. RODRIGUEZ: If I may? Because I'm confused. In the instructions for the call for measures it talks both about -- I definitely agree that this is not a health

outcome type of measure.

But the call for measures specifically refers to two groups of measures that the committee is looking for. One is the health outcomes. And within the health outcomes measure, that health must serve the health of a subset of the population.

And this is one of the areas where we consider that country of birth is an important variable to be able to identify. To observe the health of the sub-population in the U.S., many that are foreign born, because of the disparities.

And, but also, we believe that this variable complied with the second type of measure that the committee is looking for as a determinant of health.

Because again, foreign born experience a number of quantifiable socioeconomic and environmental determinants of health.

And I believe that those are very

important to, you know, to be able to collect and to identify, you know, to implement a specific intervention. And also specifically to target interventions to the population suffering specific health disparities.

So I'm not totally sure I understand. I mean, I agree that this is not a health outcome. But definitely it's a measure of this population health and disparities. And it's a determinant of health. And so I'm not sure the rationale for not accepting this variable, or this measure.

CO-CHAIR JARRIS: Mike?

MEMBER STOTO: Thanks. I mean, I wasn't able to be on the call, but I agree 100 percent with what Bobby said. I think one thing that has not been said, but needs to be said --

In fact it just answers that last question. Is that we're interested in performance measures. We often don't use the word performance. We use it in key places.

But when you look at the criteria and so one, it's performance. So which means a modifiable risk factor I think. And you can't --

So this is just like race or education or income, and so on. Clearly those are determinants of health. But you can't imagine anybody doing anything to change those things as a way of improving health.

And I think that's, you know, it's so ingrained in the health care setting it doesn't always carry over into the population health setting. But I think it's important point to make.

CO-CHAIR JARRIS: Amir?

DR. RODRIGUEZ: Mike, what's so bad -- Sorry.

CO-CHAIR JARRIS: Was it directly in response to what Mike said?

DR. RODRIGUEZ: Yes. My points that I have is how does the committee plan to identify disparities in sub-populations, all

of the sub-populations, if we don't have a state of standard, you know, we'll call it demographic or social measures, like race/ethnicity, country of birth, poverty level, et cetera?

CO-CHAIR JARRIS: You know, I think you're making a valid point. But I still agree with what Mike and Bobby said, in that we do need to understand the population being served.

And that's important for social determinants as well as other determinants of health. But those aren't -- They might be, could be very useful in sort of a risk modification.

Looking at one community versus another community. And what is the potential burden of illness in that community, based on immigration status, race, ethnicity, socioeconomic and unemployment? So we have to get our hands around that.

MEMBER STOTO: You might even want

to have different interventions depending on the nature of the community.

CO-CHAIR JARRIS: Right. But it's different from something that is a modifiable measure. For example, we certainly wouldn't want a community to say, in order to improve our health status we're going to limit the number of foreign nationals we allow in our community.

Or the opposite. We're going to increase, you know, the amount of foreign nationals, as opposed to U.S. born people of that group.

So, you know, I think that's what I'm struggling with. This is a very important thing. But it may not be a modifiable measure. Amir, please.

MEMBER QASEEM: Just to follow up. I completely agree. And that's the first thing that I was looking at in trying to figure out what does it mean.

It actually applies for 2018. And

there's some other ones as well. I just couldn't figure out what are we improving in terms of performance measures.

Second thing I was going to say was, don't we have this information available? Maybe it's just a dumb question. But U.S. census, don't they get this information in there?

Or a lot of places you can get this information? So what are we really going to add? What is CDC going to add with this measure? I just couldn't even find the --

DR. RODRIGUEZ: Yes. There is a number of, including the census, there is a number of data sources, or national data sources that collect country of birth.

But there are all the important data sources like behavioral risk factors in the system. And especially clinical kind of data, like hospital discharge data, et cetera, that do not collect that information.

And therefore, it doesn't allow us

to identify this specific vulnerable population, social population, that we could, you know, then target this specific intervention to those populations.

So I totally understand in terms of the not being a performance measure. But again, from the instructions it's not very clear to me, you know, if you're asking for a determinant of health.

And country of birth cannot by itself be modified. I totally agree on that. But it's a very good proxy for a number of legal and social and other type of barriers and exposures.

And, you know, exposure that can be very helpful in terms of, you know, targeting and identifying health disparities. Even though by itself it's not a modifiable measure.

MEMBER QASEEM: So but you're trying to get the information at the -- With the census data you can get it literally at

the zip code level. That how many folks are born outside U.S.

So why can't that information be utilized? I'm just trying to figure out why do we need a new measure? You're just trying to find out in a certain population area, how many are foreign born?

And the second question is like that one. That information, there is very good data available. I'm just trying to figure out the utility of adding another thing is there's already information already out there.

DR. RODRIGUEZ: The data is collected by the census. Again I totally agree with that, you know, how many people are foreign born in a specific area.

But the data, the country of birth is not collected in a number of key health information sources in the U.S. Like, you know BRFSS and different behavioral reflectors used in the BRFSS. Country of birth is not

collected.

So you cannot identify issues for a foreign population. If you're interested in hospital discharge data, clinical data, country of birth is not collected. So it's collected in a number of data sources in the U.S., national data sources. By half of them probably. But not collected in the other half.

And so that creates a number of gaps in our understanding of the health issues and disparities of these, you know, increasing foreign born populations in the U.S. And then it creates, you know, limitations in being able to, you know, identify those disparities and target the interventions.

We're seeing that there is a tremendous diversity within the different race/ethnic groups. And country of birth is one of the variables that can help you to understand, you know, to get more in depth into those disparities and diversity within

the race/ethnic groups.

Now I understand, you know, if the committee's looking for different kinds of measure, you know, I completely understand that. But again, I see it as a determinant of health. And so --

CO-CHAIR JARRIS: Thank you. Ron?

MEMBER BIALEK: Yes. I think the measure in and of itself is an important measure, as we talked about. And the role of NQF is endorsing existing measures, whether or not those measures can ultimately be used to make a difference, have an impact.

And I think what we're really missing here is how can one use this particular measure, which is a determinant of health, how you can use that measure to make a difference within the community in terms of health.

And I think it's important for us to note the problems that were cited with the clarity of instructions for the developer of

the measure. So unfortunately you're one of our, I guess guinea pigs here.

But it's important for us to note that clarity is an issue. And the question then becomes, you know, could this measure -- Not right now.

But could this measure be tied to specific evidence based interventions that can be impacted by knowing about this particular measure within the community? And if the answer is yes to that, that might be something worthy for further consideration.

DR. RODRIGUEZ: And the answer to that is absolutely yes. I mean, on the one side I totally agree that it's not performance measure.

But on the side, if you're asking for a measure that we can use, or how it can be used, it certainly can be used for that. It can be used to identify a specific sub-population that have the main disparities in health or specific disparities in health.

They can be extremely helpful, especially when we are in the times of limited resources, where you're going to target your interventions. And to put specific, you know, a specific population as possible to be effective.

So there is no doubt in my mind, and there is a very, very good literature right now in terms of how this variable can be used to identify those populations. And to implement intervention that are culturally and linguistically appropriate.

Again, the country of birth is a proxy for a number of additional variables that can truly help you to tune into those and to focus into those populations who have the greatest disparities.

At the same time foreign born population have very good indicators for all their measures of health. And is more, you know, the children of the foreign born, or is people who have, you know, different

race/ethnic groups born in the U.S.

Again, this variable allows you to make that differentiation, not to have this, you know, broad race/ethnic groups that we're finding that are less and less useful because they're too broad in terms of targeting public health measures. So I totally, I mean, I would say a complete yes in terms of that this measure is a useful measure for public health.

CO-CHAIR JARRIS: Thank you, Alfonso. Let's get a couple more committee members. Kurt and Matt, I think you're next.

MEMBER STIEFEL: This seems like a feature of the measures in one of the evaluation criteria that we called for, for the measures.

And we asked for a summary of measures of disparity by population group, which seems like that would be appropriate to address in the measures that were called for, as a feature of those measure. As opposed to a measure in and of itself.

In the commission paper they made the distinction of population subgroups, or sub-populations. And it seems like this is a sub-population as opposed to a measure.

DR. RODRIGUEZ: I totally agree. I completely agree. And that's probably my point from the beginning. It's not a performance measure. It's a population, sub-population, or health disparity type of measure.

But again what's not totally clear is how can we include that? This is what it sounds like. This call for measure is asking for how can we include those determinants of health in this initiative, if they are not considered performance measures? I'm not clear about that. If it's including socioeconomic, social norms --

CO-CHAIR STANGE: So this is Kurt Stange. I have a question for our NQF colleagues. There's clearly consensus that this is important to know about To stratify

other measures of health and other health determinants.

There's no question that it is a health determinant, and that it's useful thing to know. The question is whether it meets this kind of one beat performance gap thing. Because it's not, as Mike and Ron pointed out, it's not a performance measure.

So the question for the NQF colleagues, is there precedent within your NQF measures for a measure like this? Something that's useful to help interpret other measures, but maybe isn't a performance measure in itself?

And the other question is, with the disparities or any qualities group, how are they handling this? Because I imagine there would be a lot of issues like this, things that you would like to look at across many different measures of quality. Or in our case health and health determinants. And is, did we learn anything from how they've

approached this?

DR. BURSTIN: Yes. I'll speak first. And Robyn may want to weigh in on disparities. We don't have anything exactly like this.

I think the one thing that it raised for me, that was somewhat analogous, is the exercise we've recently gone through bringing in measures of research use and costs, which none of us agree in and of themselves are measures of quality. But when paired with other measures become very powerful building blocks of quality measures.

So that was the one big picture analogy I could potentially see to measures like this. As part of the disparities project, we didn't get anything along these lines that I recall. Rather more so measures that are assessing cultural competency. Not, you know, tools, patient surveys.

We did have some measures that came in on child health that perhaps get

closer. But again this issue of mutability isn't necessarily a requirement either.

There's significant variation in the population. There's differences in outcomes they have clearly demonstrated, based on the criteria we wrote in terms of outcomes for those patients. I wouldn't get as hung up on that.

But, you know, we have endorsed, for example, some of the CATI surveys that look at issues like children who attend schools perceived as safe, children who live in communities perceived as safe.

In and of themselves, the question is, what is the performance measure. I guess that is the performance measure of the community.

And I guess I'd be curious about, you know, Paul and Bobby's perspective on a, you know, how in some ways are a measure like that inherently different than a measure like this. But Robyn, anything from the

disparities perspective that you want to add?

MS. NISHIMI: No, we didn't, as you said, get any measures along this line at all. And so then the focus was, in that project, what elements do you use to stratify reporting for other measures?

CO-CHAIR JARRIS: It was my understanding when we discussed that health disparities group though, that in fact, it was limited to health care, within the health care setting. So a measure like this wouldn't enter.

DR. BURSTIN: The call for measures wasn't limited. All we got in was health care. So we could have gotten something broader, it just didn't come in.

MEMBER STOTO: I wanted to second the point that the developer and Ron made about the documentation. There's two things. One is that the current criteria and the form for filling it out are the ones that NQF has always used for its health care measures.

And really hasn't been modified to represent the population health. We put the change in the criteria, but the form hasn't been modified. And I think that's an important issue that has to be addressed.

And I also think that the background paper doesn't do a good enough job at this. And we probably want to address that too. And I think these are really issues for tomorrow's discussion. I think they really have a lot to do with the relatively small number of measures that we've gotten in moving forward.

MEMBER KINSINGER: Hi. This is Linda. If I could jump in? It's hard being on the phone, not being able to raise my hand. I just, I also reviewed this measure, and want to concur with everything that's being, that's been said already.

That it doesn't measure a gap in quality performance. The information that was presented spoke eloquently towards the reason

why place of birth is, you know, leads to disparities.

But didn't talk about what's currently measured, how this could help inform interventions that could address that. So just to concur with everything that's been said.

MEMBER PESTRONK: Now Helen, just in thinking about the example that you gave, perceived safety, children's perceived safety. There one can clearly see how that kind of measure could lead to performance, on the face of it.

Because levels of perceived safety could change. And activities or steps taken to decrease, increase perceived safety in this case, would be a change in performance in the direction desired. And one would be able to determine whether the activities themselves, in fact, produce that change in performance.

And I think that's the distinction that we're trying to make here. Just on the

face of it, knowing the place of birth tells one nothing.

Or it doesn't tell one nothing, because I think the explanation that has been provided about it suggests how it might be used. On the face of it, it isn't a measure of performance.

CO-CHAIR JARRIS: Can I make a suggestion? I'm wondering if this very important measure is a denominator, as opposed to a complete measure.

So for example, if we looked at foreign borns in a community, and the numerator was something like the percentage of foreign borns with early access to prenatal care, with access to a medical home, with complete vaccinations, with high school graduation rate within three years of ninth grade.

Then we would have a real population health measure. But I think we're talking now about a denominator. And a very

important one.

MEMBER PESTRONK: Yes. Be attached to something else.

MEMBER STOTO: A strata I think. Just like race, ethnicity, education, income, all those things.

MEMBER BIALEK: This may be something for tomorrow's discussion. But I pulled up the HHS consensus statement on quality in public health. And I was looking at the aims there.

And I'm thinking that if developers of population measures consider these aims, you know, is this measure population centered? Yes, no. Does it work toward achieving health equity? Yes, no.

Can it be used for formulating policies, you know? Yes, no, and how for each of these. And that could be quite helpful I think for developers of measures to note the population health impact. And, you know --

DR. RODRIGUEZ: And again, for all

those questions my answer would be yes. So this is why I'm a little bit confused. Because it's not a performance measure, but it seems it fulfills many of the criteria, or the questions that you are asking.

The question will be yes. It will be, as variable it will be very useful to the measure. And let me just give a very, very simple example as in racial discrimination. There is quite a lot, as you know, in terms of discrimination in access to care and quality of care.

And one of the criteria for discrimination is discrimination to national origin. And how, so how can you monitor discrimination and access to care in the U.S. related to national origin if you don't collect country of birth.

That's one of the most critical, I believe, uses of this variable. You cannot identify that kind of discrimination based on, or different quality of care based on

race/ethnicity or anything else. Or any performance measure that you're looking at would allow you to look at that important determinant of health.

And the same thing, the quality of care in health insurance. Again, country of birth allows you to identify differences in genetic background, behaviors, legal access to care, language issues, cultural issues.

Many of those are modifiable. They country of birth by itself is not modifiable. But many of the factors associated with being born in another country and living in the U.S. are modifiable.

So this is where we think it's a very useful measure for, you know, policy, interventions, identifiable collections, and monitoring disparities and reducing disparity gaps in the U.S.

CO-CHAIR JARRIS: Okay. Why don't we -- Matt and Bobby, and anyone who wants to speak after that has to buy us breakfast on

Friday. Because we'll still be here.

MEMBER STIEFEL: Helen, you mentioned impact is not a criterion. But I --

DR. BURSTIN: I said the fact that it's mutable isn't necessary. Right.

MEMBER STIEFEL: But that a crux of this issue I think. And it's I think quite related to impact. It seems like this one doesn't meet the impact criterion, the first and fundamental criterion, in and of itself. It's important. But, and so it's highly related to mutability --

DR. RODRIGUEZ: If I could --

MEMBER STIEFEL: So it's an important general point I think, that we'll probably come back to for a lot of these measures.

DR. RODRIGUEZ: Sure. Especially impact.

DR. BURSTIN: Yes. And as you just -- It's interesting. I feel like this is just an important exercise for the next

measure as well. So I think it would be good to just kind of get it done.

But we do specifically in our population health guidance here say, demonstrate a high impact of health effects. A large population who has a substantial impact for a smaller population is a source of significant health disparities, leading cause of mortality, morbidity, et cetera.

So I mean, it certainly seems like it's potentially a source of healthcare, significant healthcare disparity. So to me it's not an issue of impact. I think it's more a question of, I think it's a building block of a measure.

And the question is, is it on its own standalone? Or is it really something that's a building block? And I guess the question is, since we did specifically call for social determinants, I think it's helpful for us to at least think through what we meant by that.

And maybe that's more for tomorrow. Because certainly this is a social determinant of health, which has a significant source, you know, that leads to significant health disparities.

Certainly somebody who practiced in a, you know, refugee clinic for ten years, this is a huge issue. I just think it's important for us to at least think it through. Because we did specifically indicate that would be high impact.

It is a source of disparities. It is a social determinant of health. I'm not sure exactly how it would be used. But I think that's been a question all along.

DR. RODRIGUEZ: But again, that's what I'm a little bit confused. You're saying that you're now asking for social determinants. But in the call for measures it says that you're asking for health outcomes and determinants of health.

So you're asking for it. So

that's where I'm really confused. If you were telling me we only looking for performance or health outcome measures, I would totally agree with you that this is not a pertinent measure.

But if you were asking for health outcomes and determinants of health -- And like everybody seems to be agreeing that this is a determinant of health, then I think it's totally relevant. So that's what I am really confused with the rationale.

CO-CHAIR JARRIS: So I think we want to thank the developer for helping us to clarify the call and the criteria here. Just as a footnote, if anybody read the newspaper this morning. There is some question about the mutability of place of birth. But I think that's a specific case that relates to the political context.

DR. RODRIGUEZ: Yes. So that would very helpful clarification. And then so that the determinant of health are included in these call for measures, or not. And if yes,

you know, how you compose determinants of health, without being a performance measure.

CO-CHAIR JARRIS: Alfonso, I want to thank you. This has been very helpful. And as you can tell, we're really in still a developmental phase here. And your confusion is something we created inadvertently. Because we're all confused ourselves.

And what you put forth is a very important measure. The question is, you know, and we've struggled with this about the NQF criteria. And does your important measure fit into this NQF process? Or would it be considered a building block within this process?

But I'd like to -- I think to punish you, we should put you on the committee for doing good work. But thank you for what you've done. And I hope you continue to work on this. And we need to go now probably to the voting process and move along. All right. You're buying lunch?

MEMBER SPANGLER: I had a question you may want to answer. I mean, he asked about what social determinants. Should we have social determinants? And I think we kind of answered that.

I don't know if he heard, but there are social determinants that are modifiable, and that we can try to improve on. This just isn't one of those. So maybe -- I don't know if that answers your question, Alfonso.

CO-CHAIR JARRIS: It's not one we'd want to modify. And I have to say I have a little sensitivity to that, given some of the anti-immigrant sentiment in this country.

And parts of it I would want to be very careful about putting place of origin up as some kind of a quality issue. That's why I would get back to, this is an important denominator. And we need a numerator on top of it.

DR. RODRIGUEZ: Okay. Thank you

very much.

CO-CHAIR JARRIS: So why don't we turn it over to Elisa and Kristin to go through the voting process?

MS. MUNTHALI: Okay. Thank you. I just wanted to remind everyone that importance to measure and report is a threshold criterion that must be met in order to recommend a measure for endorsement.

And so all three sub-criteria must meet, must be met to pass this criterion. So the first one is 1A, impact. And I don't know if you can see it.

Addresses a specific national health goal priority or data demonstrated a high impact aspect of health care, numbers affected, resource use, severity and consequences.

And so the rating scale is one for high, two for medium, three for low, four for insufficient evidence. And so Kristin will give you 20 seconds to rate 1A impact. And

make sure you're pointing towards this monitor.

I think two people or three people we still haven't received. So if you can just -- Okay, we've got two more. One more. I think we have nine out of ten. And Sarah and Linda, if you can tell us your votes? High, moderate, low or insufficient evidence.

MEMBER KINSINGER: This is Linda. I would say insufficient.

MEMBER LINDE-FEUCHT: This is Sarah. I agree, insufficient.

MS. MUNTHALI: Okay. So we have two -- I think we have four for moderate, five low, and 3 insufficient. Okay. So we move on to 1B.

(Off microphone comment)

MS. MUNTHALI: I don't know if they abstained or -- Okay. Make sure -- Would you like to -- We can do it again. And make sure you're pointing to this monitor.

Okay. So that's it. Okay, we

have ten responses. Okay. Zero for high, three for moderate, six low, and three insufficient. Okay.

DR. BURSTIN: Basically this is a must pass criterion. So you must pass all three to meet it. So the first sub-criterion is already not met. We could finish importance if you like. But essentially --

If you think it would be a useful exercise, just for the discussion tomorrow, to actually see which of these sub-criteria are not being met, it might be useful to continue the next two. But, okay.

MS. MUNTHALI: So to 1B? Okay. So this is performance gap. Data demonstrated considerable variation or overall less than optimal performance of cost providers and/or population groups in disparities in care. And so same rating scale.

And I think Kristin will give us 20 seconds again. I think we have four people, one person missing. You can just step

-- That's it. Okay, Sarah and Linda? For performance gap, 1B.

MEMBER KINSINGER: Insufficient.

MEMBER LINDE-FEUCHT: This is Sarah. I agree.

MS. MUNTHALI: Okay. So we have zero high, one moderate, six low and five insufficient. Okay. We'll go to the final criterion, for importance to measure and report. It's 1C.

And this is evidence. Health outcome with rationale. Or quantity, quality, consistency of body of evidence are met as follows. Consistency, moderate or high. And quantity and quality, moderate or low. Or low with special circumstances.

And this is one for yes, two for no, three for insufficient evidence. Okay. Six missing in the room. I think we have two missing in the room now. And we have all ten in the room. And Sarah and Linda?

MEMBER KINSINGER: Linda. And

again, I think it's insufficient.

MS. MUNTHALI: Okay.

MEMBER LINDE-FEUCHT: Sarah. Insufficient.

MS. MUNTHALI: Okay. So we have one for yes, four no's, and seven insufficient. And so the measure doesn't pass. And we stop voting.

MEMBER STOTO: Can I just flag an issue for later? I mean, there are two things here really. One is, when we say the evidence is sufficient or not, you know.

In the healthcare setting we mean evidence that doing this process will lead to the outcome that we want. It's not clear what the appropriate evidence is here.

If you thought, what's the evidence that place of birth is related to outcome? Well that evidence is strong. What's missing I think, in my view, is doing anything --

Because we can't imagine what to

do, is there. So I think it's something that's implicit in a known process. But needs to be more explicit here.

CO-CHAIR JARRIS: Yes. So the evidence should be that whatever is modified or when an intervention occurs, actually improves health.

MEMBER STOTO: Yes.

MS. MUNTHALI: Okay. This is also from the CDC, the same developer, Alfonso. He introduced this earlier. This is measure 2018, the year of arrival to the U.S. for foreign born nationals. Alfonso, or -- I think, Alfonso, did you have anything else to add to your introduction?

CO-CHAIR STANGE: I guess, since many of the issues are the same, focusing on what's new in relation to the previous measure.

DR. RODRIGUEZ: Yes, sorry. It's a related variable, so a lot of the things have been said about the previous variable are

relevant to this one.

It allows you to identify, you know, recent arrivals versus people who have been longer in the U.S. And again, the purpose is to show that foreign born have been longer in the U.S., their health gets worse for a number of health indicators.

And again, allows you the opportunity to identify new arrivals and the opportunity to implement prevention services and activities to, strategies to produce, to prevent those, you know, health indicators for becoming, getting worse.

Again, in terms of the, what can you do with this information. If you know the specific countries of birth, people from the countries have experienced, they experience is in disparities in access to care or in specific health outcomes.

Then it allows you to use the census or other health information to, or information to know where those people are,

where the languages is, other issues related to legal access to care, to culture, to behaviors, and then to implement specific intervention.

So this is what we think this variables are important to identify those disparities, and to target the populations to review. So that you have the disparities, without these variables it's very difficult to do it.

CO-CHAIR JARRIS: Thanks. So I was one of the folks that led to some of the, who participated in this phone conversations. And I think, rather than repeating much of what we discussed, again, this could be an important measure for a community.

But if it were -- To make it actually an intervention or something that made it a quality primer, we would need it to be paired with something that you did.

So for example, again I go back to it will be very important to measure the rate

of enrollment in early pre-natal care among this population.

One community compared to another as an indicator of how they are welcoming, how they are reaching out to this population. But just knowing the demographic characteristic itself doesn't really identify or address the performance gap.

DR. RODRIGUEZ: But I would make the case that only having the quality measure of the performance measure without having these variable doesn't allow you either to put aside these disparities.

So again, I really appreciate the input from the committee. It's very helpful. And I would be very interested in further exploring with you, you know, how can we incorporate some of these --

Not only the country of birth, but all of the variables into these call for measures. Because this population -- The measure, the performance measure of the

quality measure by themself aren't going to allow you to do that. You need determinantal health measures too. And this is one of those.

So how can we incorporate those measures within this call for measures? And I would be very interested in, you know, following up conversations. And looking for opportunities to include those type of social determinants.

CO-CHAIR JARRIS: It's an important point. And I ask the NQF staff. Is there ever a process in which you, if you will, endorse a measure around the social determinants of demographics of a population as a building block, as opposed to --

Because it's true. If you had this measure you could put something on top of it, multiple interventions that would be important. We don't have those interventions in this proposal right now.

But is it worth taking, looking at

a two step process in which we say, okay, here is a standardized, valid way of looking at year of arrival to the U.S., number of foreign borns, you name it.

Children living in poverty in this community, or whatever it happens to be. And then put that out there for other measure developers to put on, you know, interventions on top of them.

DR. BURSTIN: Never been done before. But then again neither has population health. So, you know, the inclusion of social determinants of health on the call is something that I think potentially, as we discussed it up front, might be used in that way.

DR. RODRIGUEZ: But again, to do it is very confusing. Because the call for measures, I'm looking at that right now, talks, asks about both, health outcomes and determinants of health. So those I am confused about.

But again, I will be very happy to, you know, to continue the discussion and to look and see if there's any opportunities to include some of these measures, and how to do it.

Because I totally agree that, you know, the background or the rationale that is being asked doesn't allow to enter this kind of determinant of health. Even though in the call for measures, it's asking for them.

DR. BURSTIN: The call did say modifiable socio-economic and environmental determinants. And perhaps the sticking point we're having here.

DR. RODRIGUEZ: But it talks about -- It talks also about genetics. And it talks about other -- So again, it's sort of confusing.

CO-CHAIR JARRIS: Okay. Let's go to Ron, Matt, Bobby, Madeline.

MEMBER BIALEK: I think the idea of pairing a measure is an excellent one. And

I also think that there's another option, which is for the developer of the measure to demonstrate, with evidence, that knowledge of that measure, knowledge of that information leads to improvement in health in the population.

And you can do that. I mean, the guide to community preventive services uses that type of an approach. In this instance, again, it's a high impact for a population. But it's a low impact in terms of any evidence related to how one can make an improvement.

So again, a topic for tomorrow's discussion is, if we deal with the social determinant, does it have to be paired? Or can a developer demonstrate evidence that knowledge of that determinant can lead to a change that improves health?

MEMBER STIEFEL: Just to make the point that this is an evaluation criterion in the call for measures. And evidence of disparities by population group

So a good measure, if there are significant disparities associated with place of birth or date entered the U.S., then that should be one of the components, one of the strata of the measures that we endorse. That was quite explicitly a part of our evaluation criteria.

DR. RODRIGUEZ: But then again, it sounds good to explore it again. Because it's defiantly a strata that should be collected. And there's plenty of information in the literature that shows the disparities by country of birth, after adjusting for many other, like gender, poverty, health insurance, many, many other variables.

Still the disparity on country of birth persists. Even within the same race/ethnic groups. Hispanics foreign born versus Hispanics U.S. born.

MEMBER STIEFEL: So an example of the measure. Say, take BMI. If place of birth or date entered the U.S. is an important

stratum for BMI, then that should be an element of the BMI measure.

DR. RODRIGUEZ: And the answer is yes. The answer is a complete yes. This is a complete yes in health.

MEMBER STIEFEL: But then the measure is BMI. And this is a stratum of the BMI measure.

CO-CHAIR JARRIS: So in other words, would it be BMI as the numerator and among this population? And the total of this population as a denominator?

MEMBER STIEFEL: It would be BMI as the measure. And then BMI would be stratified by various important sub-population.

DR. RODRIGUEZ: Yes. Completely agree on that.

CO-CHAIR JARRIS: Okay, Alfonso, let's let the other committee members speak. And then we'll get, let you have a last word before we vote. I think I had --

MEMBER PESTRONK: I like the fact that this is pushing the discussion at NQF to consider things that it hasn't considered before. Because that's been the challenge for the population health committee to begin with, to introduce a different way of thinking about measures.

I think that in the discussion tomorrow, one way to perhaps address this is to think about whether a separate evaluation sheet for a proposed measure might be proposed, to gauge the potential acceptability for stratification type measures.

Rather than looking at them as having to have the same kinds of characteristicness as this worksheet requires them to have.

Because it might be, it could be useful down the road to have NQF endorsing characteristics, like these two that we've just considered, as a way to point out their importance as means to stratify.

And perhaps even to say that these are ways to pair traditional NQF measures with other acceptable, with acceptable means of stratifying those traditional measures.

Because it could push the issue into the other domains where NQF has measures also, where these stratification elements should be considered, but aren't necessarily considered.

CO-CHAIR STANGE: I'm interested in having this discussion go forward beyond the life of this committee, which might -- We'll see what our charge is tomorrow. But we might be ending.

And I think that these two measures are relevant, not just to population health and population measures, but to clinical measures as well. And they're quite actionable.

I mean, a number of European countries, for example, have measures of social deprivation. And they pay differently

based on the social deprivation index.

And certainly everybody's talked repeatedly about how these can be used to stratify. And I think that's true for both clinical and population measures.

So if we end up not going forward as a group, but the disparities group is going forward, I think this would be a wonderful thing to ask them to take up. Because that would be a way of really amplifying the impact of the disparities group.

That they're not just looking narrowly at measures of disparity and equalities, or equity. But that they're actually looking at how disparities can be looked at across all the NQF measures.

MEMBER NAEGLE: Yes. And in that way I just wanted to comment that I'm very glad that we've moved in this direction. Both in relation to thinking about disparities and looking at the social determinants.

But I feel that in looking at

attributes or variables, it is so particulate that it's not meaningful in terms of our going forward and looking at measures.

So a number of important points have been made about embedding those either as strata measures, or creating a link between those particular indicators and some types of intervention.

And while we have quite a lot of evidence about interventions related to social determinants, that evidence is very heterogenous.

And I think we are at the very beginning of trying to figure out how to use things like social determinants in creating measures that are meaningful for performance. But I'm glad we're moving that way.

CO-CHAIR JARRIS: Yes. So the question will be, and Kurt, you raised this issue about what is the future of this committee? And is tomorrow our last day or not? But this issue can't be lost.

And in Maryland, not far from here, they took on a very important infant mortality initiative. Set a goal to lower infant mortality by ten percent within two years. And they hit it slightly about, within 18 months into it.

But then they re-looked at their data. And they had done nothing to close the gap between African American and Caucasian infant mortality. In fact, the decrease had been in the Caucasian population. They increased the gap.

They came back and set a specific measure for African American infant mortality, and actually hit it. And started to close the gap. So I think that's the importance of these sub measures, is to say okay, we've just improved the health of this community.

But actually we did nothing for the foreign borns or the people that have been -- So we need to re-look at what we're doing. So somehow we have to capture this an not lose

it.

DR. BURSTIN: Just to follow up on that. Part of the work the disparities committee has been doing has been updating our criteria for what is disparity sensitive performance measure.

And there is a set of criteria about the prevalence of the condition in a given population, the size of the quality gap, evidence of what could be done to improve the situation.

So we actually have a team that's been going through the entire portfolio particularly focused on outcome measures. Indicating which one of them, including infant mortality actually, which should always be stratified explicitly, because we know there are known disparities in those areas.

So maybe we'll try this. We'll actually make sure we distribute that, those criteria in the morning, just to help this discussion further.

CO-CHAIR JARRIS: Are those people -- Could they come and speak to us for a few minutes tomorrow? Could we get them in here for half an hour?

DR. BURSTIN: Robyn's the consultant lead on it. She could certainly do that, yes. On the disparities.

MEMBER KINSINGER: Hi. This is Linda. If I could interrupt to make a comment. I'd be interested in knowing the reliability of reporting year of entry into the U.S. No information about that was presented.

And I just wonder -- There are, I think, people who come in and out. And so what year would be the right year to report? Are there reasons why people might report something, you know, to gain some advantage somehow?

I don't know what that would be, you know. Just wondering whether there's, whether this is actually even as reliable a

measure --

Certainly place of birth seems pretty clear to me. But year of entry seems like one that might be hard to be certain about its reliability and validity.

DR. RODRIGUEZ: If I may respond directly. It's a more difficult measure to obtain. The developers have done a lot of work in terms of how to ask that.

I mean the way that we're proposing is actually to use the census wording, where they have pointed when you came to live in the U.S., not when you came to visit or to stay for a period of time. When you came to live.

It's still very some limitations in the data. Especially when you have, you know, anti-immigrant, you know situation, political situation. It might be difficult to get.

But it still doesn't mean that it's not important to try to obtain. We know

that we have a lot of issues with the race/ethnicity, you know, quality of the data and the completeness.

That doesn't mean that it's not important to collect. But there are challenges in collecting this variable compared to the country of birth. So I agree on that.

CO-CHAIR JARRIS: Matt, and then Madeline.

MEMBER STIEFEL: Paul, to your question. There's a good example I think from the U.K., the NHS. They don't have a life expectancy goal.

But they do have a goal for the disparity between the first and fifth quintiles of the deprivation index for life expectancy. So I think that's the kind of thing that we're talking about.

CO-CHAIR JARRIS: So somewhere in there, there were building blocks for deprivation index that had to get started.

And I think that's our question. Do we look at building blocks.

MEMBER NAEGLE: Yes. I wanted to speak just a little bit to the comments you made in your rationale about health literacy. I'm just remembering that health literacy among our mainstream population in this country is relatively low.

So when you're trying to collect this data, which we can see is relevant to health status, explaining that to immigrant populations, or having immigrant populations understand the meaningfulness, or the usability of this kind of information, it would seem to me would even complicate our collection of it.

And I think that's something that CDC may want to think about in terms of how we gather information which we see is very relevant to the development of both measures and targeting interventions. But really is not a wide -- Well we're going to have some

difficulties getting reporting from the public, for sure.

CO-CHAIR JARRIS: Any of the committee members on the phone have a comment? And if not, why don't we move on to the voting process. Okay. Elisa?

MS. MUNTHALI: Okay. We will start again with 1A, impact. The rating scale. One high, two moderate, three low, four insufficient. And I think time has already started. Six people we're waiting for in the room. One more person. Thank you. Sarah, Linda on the phone?

MEMBER KINSINGER: I think I'd say insufficient.

MS. MUNTHALI: Okay.

MEMBER LINDE-FEUCHT: This is Sarah. I'd say low.

MS. MUNTHALI: Okay. So we have three moderate, we have seven low and two insufficient. For 1B, opportunity for improvement. Performance gap. Sorry. Same

rating scale. And time has started. One more person. Thank you. Linda and Sarah?

MEMBER KINSINGER: Low.

MS. MUNTHALI: Okay.

MEMBER LINDE-FEUCHT: And this is Sarah. Low.

MS. MUNTHALI: Okay. So we have one moderate, ten low, one insufficient. And for 1C, evidence. This is one for yes, two for no, three for insufficient evidence. And time has started. Four more people in the room we're waiting for. Three more. Two more.

I think one more. Okay, I think that was nine, right? No, ten. We got ten. Linda and Sarah?

MEMBER KINSINGER: No.

MEMBER LINDE-FEUCHT: And this is Sarah. No.

MS. MUNTHALI: So ten no's and two insufficient. So the measure doesn't pass.

CO-CHAIR JARRIS: Well this was a

extremely helpful conversation. And I think it's going to be very informative to this process.

So although the measures did not make it through the criteria as they exist today, this has been a very valuable process. And we want to thank you for the opportunity to have this discussion, and really move our thinking along significantly.

And we're looking forward to figuring this out, so these measures can come back to us in a way that meets the meaningful, criteria set in a meaningful way. So thank you.

DR. RODRIGUEZ: Well thank you very much. I appreciate the comments from the committee. I look forward to working with you in the future.

CO-CHAIR JARRIS: So I believe we actually are -- We can take a break now, a little ahead of schedule. And then we can come back in 15 minutes. So about 25 of the

hour. Does someone on the phone want to say something?

MEMBER LINDE-FEUCHT: Yes. Paul, this is Sarah Linde-Feucht. I actually have to sign off for a little bit. I am not positive I'm going to be able to rejoin today.

I do have one of my HRSA colleagues, Ian Corbridge, on the phone. I don't know if the procedures of NQF allow me to give him my proxy vote in my absence.

But I just -- If so that would be great. If not, I just wanted to say, if I'm not answering, it's because I'm not on the phone.

CO-CHAIR JARRIS: I suspect that that, what we probably need to do is we will send out a posse to find you if we fall below a quorum. But otherwise we'll continue.

MEMBER LINDE-FEUCHT: Okay.

CO-CHAIR JARRIS: And hopefully you'll be able to join us tomorrow. Because I'd be very interested in your perspective

from, to the extent you can reflect a broader HRSA perspective on the measures. And why we might or might not have gotten more measures from HRSA.

MEMBER LINDE-FEUCHT: Yes, okay.

CO-CHAIR JARRIS: Thanks. Break time?

MS. MUNTHALI: Anika?

(Whereupon, the above-entitled matter went off the record at 10:21 a.m. and resumed at 10:40 a.m.)

CO-CHAIR JARRIS: Okay. So the next measure we'll be discussing is 1999, late HIV diagnosis, and 2020, adult current smoking prevalence. And, Pascale, are you on the line?

DR. WORTLEY: I am. Irene Hall is -- I haven't been able to reach her. So she may be just joining me at 11:00 a.m.

CO-CHAIR JARRIS: Very good. Would you like to get started? Are you able to?

DR. WORTLEY: You know, well depending on the nature of the questions, it probably would be good to have her here, since she's, you know, more expert in the surveillance system itself than I am. So it kind of depends on the nature of the questions.

CO-CHAIR JARRIS: Would you be able to do an introduction? Or should we go to the committee people who reviewed this?

DR. WORTLEY: An introduction as in describing the measure?

CO-CHAIR JARRIS: Yes.

DR. WORTLEY: Yes, I could. I wasn't really sure, you know. I didn't really know how this was going to play out. But I'd be happy to.

CO-CHAIR JARRIS: We're making it up as we go.

DR. WORTLEY: Yes. I can do that.

CO-CHAIR JARRIS: Please do.

DR. WORTLEY: Oh, okay. I wasn't

sure if everybody could hear me at this point. So I'm Pascale Wortley. And I work in the division of HIV AIDS prevention.

And my colleague, Irene Hall, who heads up the surveillance branch, will be joining me shortly. So we proposed this measure on late testing because it's something that we would like to see change over time.

As described in the document that you received, the late testers are those who each year are diagnosed with HIV who are essentially diagnosed at the same time with Stage III HIV infection, that is, AIDS. Generally based on a CD4 count under 200.

So obviously at this point in the epidemic we would like to not be seeing such a substantial chunk of people diagnosed late in their course. And therefore not having been tested earlier.

We are holding steady at around just a little under 30 percent of new HIV diagnoses every year, representing people who

are having a concomitant Stage III diagnosis.

We expect to see changes in this measure as testing rates increase. Or as the right people are tested. And I'm saying that because routine opt out testing is recommended for facilities where there is an undiagnosed prevalence in HIV greater than 0.1 percent.

But that obviously doesn't apply to everywhere in the country. But in lower prevalence areas there's more of a target, focus on targeted testing, that is, reaching high risk people.

In any event, the result of effectively applying these approaches should result in a decrease in persons diagnosed late in their course of illness. Does that cover what you need out of an overview?

CO-CHAIR JARRIS: Yes. Thank you. And Mike or Linda, did you want to speak to this?

MEMBER STOTO: Can I do the -- Hi Pascale. This is Mike Stoto. We worked

together on --

DR. WORTLEY: Yes.

MEMBER STOTO: -- unrelated issues --

DR. WORTLEY: I remember you.

MEMBER STOTO: -- some time ago. And I think this a really very interesting proposal. And Pascale, in the earlier discussions this morning we were confronting some difficulties, I think with the documentation with which we ask you to justify these things.

And I think this one presents a different set of problems on our end in what we're asking for. I think that having worked in this, in the field of HIV for a long time, to me this actually makes a lot of sense.

Because what it really is, is a way to measure whether or not people have been tested, screened and tested early. It sounds of course there is that you don't know how many people are HIV positive for the

denominator.

So what they do, like is done in cancer, they look for people who are diagnosed at a late stage, which is what this is all about.

DR. WORTLEY: Right.

MEMBER STOTO: And so the intervention that really is the focus of this is screening. Spreading the spread and utilization of screening.

So the evidence, questions about evidence really ought to be about, what is the evidence that screening makes a difference? Both to the individuals themselves, in terms of their outcomes. And in terms of preventing spread to others.

And I think that the evidence there is actually quite strong, and has been for some time. And of course there are strong recommendations from CDC about universal screening and things like that.

So I think that if understood that

way, the evidence for the importance and the impact of this is quite high. But I think that because it really hasn't been written up quite in that way, I think that presents a bit of a difficulty to the committee.

DR. WORTLEY: Oh, in terms of what we submitted.

MEMBER STOTO: That's right. And I think --

DR. WORTLEY: Yes. And, you know, to tell you the truth I felt like I was trying to fit a little bit a square peg in a round hole.

MEMBER STOTO: Well you were. You were. And for a number of reasons. And I think that this is, like I said, this is something that we have to do at our end, is to improve that documentation request process.

But I think that understood that way -- And also the other thing I should say is about disparities. I mean, they do, the documentation does make a good case for

disparities, a number of dimensions, performance gaps, and so on.

So I think that if understood that way, the evidence is quite strong for this. So I think it's a problem of documentation, not of the merit of the measure.

The other thing I would say is about the scientific properties, validity and reliability, and so on. And here, and maybe your colleague once she gets here, will have something more to say about this.

But the, what the documentation basically says is, this is all in the surveillance system. And what it doesn't address is, how does it get into the surveillance system?

And, you know, there are issues that I know have been thought, people have spent a lot of time thinking very carefully about.

Like suppose somebody is first diagnosed in Arlington, Virginia. And then

has an AIDS diagnosis in a hospital near where I work, from which you can see Arlington, Virginia, but it's in D.C. Does anybody really know that people are matched up the right way, and so on. That's critical for getting this right.

DR. WORTLEY: Right.

MEMBER STOTO: And I believe that CDC has worked hard to get that as good as possible. But I think that's the kind of thing that we need to be thinking about here.

DR. WORTLEY: Yes. And Irene discussed that on the last call. And I think the example you're picking is the perfect example of a tangled area. So, you know, she can address that a lot better than I can.

I do tend to think that even though on the last call we discussed using this measure at different levels, because of those challenges and some other reasons, I think it's best used at a more macro level, like a State, than trying to dive deeper.

MEMBER STOTO: Which is I think fine for the population that we're talking about. I think that's the appropriate level too.

CO-CHAIR JARRIS: Can I ask, could it be if the -- Is it possible to actually look at the point of testing? Although -- And look backwards to say, is there evidence or documentation that there was a referral from, or a followup and a referral on the test from the point of testing? I know, I think --

DR. WORTLEY: Well that would be hard through the surveillance system. Probably not really doable. It would require different data sources.

CO-CHAIR JARRIS: Linda, did you want to add anything.

MEMBER KINSINGER: Yes. So I was -- HIV's not an area that I work in a lot. So I may be coming at this not as well informed as others.

But I couldn't understand the

rationale for changing the time frame from 12 months to three months, when in fact, the documentation that was submitted says, almost 85 percent of persons diagnosed with Stage III within 12 months of a diagnosis were in fact diagnosed within three months.

So it seems like most everybody's giving diagnosis in three months. But I couldn't really tell from the justification that was given how changing the measure from 12 to three months would improve reporting. And perhaps would drive screening. I just couldn't make those links together. So that was my concern.

DR. WORTLEY: let me try to clarify --

MEMBER STOTO: Can I just add one more thing to that? I think that you might argue that in fact many people -- I'm sure it's true that many people are diagnosed with AIDS at the same time they're diagnosed with HIV. Not just three months or 12 months, but

simultaneously. So I wonder even why it should be not zero, rather than --

DR. WORTLEY: Right. So the reason -- Let me address that first. The reason it's not zero is because we need to allow some time for that initial CD4 count report to come in. And somebody might --

Imagine somebody gets tested, gets their test result today. And then they're going to go see a provider. And that provider will do a CD4 count, and lo and behold, it's 190.

But that doesn't happen the same day. But we would expect that it would happen within three months. So that's the reason for that. But they are in fact -- Mike, you're right. At the time that person was tested they had AIDS, they just didn't have the documentation of it.

MEMBER STOTO: So then moving to 12 months would --

DR. WORTLEY: But the 12 months is

what we're moving away from. So that is an old measure that we're moving away from. And if Irene is on the line now --

DR. HALL: Yes. I'm on.

DR. WORTLEY: Oh, she is. She may want to address why we used to look at the 12 months and are now moving to the three months. And just to clarify again, the reason we didn't provide data for the three months, even though that's the new thing, is because that data is going to be released soon.

And we can't be releasing here first. But I'm going to turn it over to Irene. Because she's the best person to address that.

DR. HALL: Yes. The problem with the 12 month measure is that it really is a combination of late diagnosis and if people get treatment in time.

So it's really not a good measure to determine what the stage of disease is at diagnosis. So we are moving now to a measure

that we call stage of disease at diagnosis, which is in line with our case definition. And as --

I just heard a snippet here from what Pascale was saying, that we define it as a measure that we received in three months, because the person gets diagnosed and they need the care. And we need to get the lab report in a certain --

We need to allow a certain amount of time for people to go into care and get that first CD4 measure. And we just define that as three months.

I mean, in theory everybody should be seen by care within 30 days. But as you know, that doesn't always happen. So that's our reason.

DR. WORTLEY: So it is meant to be a measure of concomitant, of being Stage III at diagnosis.

MEMBER STOTO: That makes a lot of sense to me. Irene, I don't know if you heard

the earlier question about matching up people who may have been diagnosed with HIV in one State, with AIDS in another State, and so on.

DR. HALL: Yes. So we are -- Every six months we send out potential duplicates that we find in a national data base. And those thousands of those potential duplicates, they go out to the State partners. And they resolve whether these people are the same as, or different than.

And they mark that in the database, so that we at the national level know which reports we need to combine as one report, versus which reports are actually two different persons.

MEMBER STOTO: Okay. I thought that was done. Glad to hear that. Thank you.

CO-CHAIR JARRIS: Linda, was your question addressed?

MEMBER KINSINGER: Yes, I think so. Thanks.

CO-CHAIR JARRIS: Kurt.

CO-CHAIR STANGE: So it seems to me that it is a population health measure. And that interventions by probably multiple stakeholders would primarily affect the denominator. And the effect would be to access to care.

And then what happens in care, as far as how sensitive that is to early detection. But also to screening programs. There could be population or community based screening programs. And then maybe indirectly by mechanism --

I can't really quite do the logic diagram in my head. Through prevention programs. But it does seem like it has multiple determinants that are more than just what happens inside health care. So it does seem like a population health measure.

MEMBER STOTO: Well counseling and testing is one of the main HIV prevention programs.

CO-CHAIR STANGE: Right.

CO-CHAIR JARRIS: Other comments, questions? Amir?

MEMBER QASEEM: How is this measure going to take into account home measurement that's going to be coming?

DR. WORTLEY: The home test?

MEMBER QASEEM: Yes.

MEMBER QASEEM: You know, we don't know. I mean, at this point we really don't know what the uptake of the home test will be. The idea is, is that somebody who was to test positive at home would then present for care. And then that provider would test them again.

CO-CHAIR JARRIS: So could I hypothesize on that? Because this could very well be a measure of what happens in the clinical setting. But if we -- And so a clinical measure.

But if we go to our second and third sort of tiers, the behavior, healthy behavior and lifestyles, they're messaging around safe sex. But messaging around early

testing, you know, before you get engaged with someone, make sure you're both tested.

All that kind of messaging can affect this also. So as well as -- So maybe the partners choose to home test. But if you have an environment in which that means, I better go get evaluated right away, and you've created an environment where people can do that in a welcoming and safe way, it will happen.

So I think home testing should feed people more in a early fashion into clinical testing, you know, diagnostic testing in a clinical setting. So what I like about this is it affects the clinical setting. It affects the healthy behaviors.

And also you can look at policy. Do you have needle exchange programs that are actively doing prevention counseling? That are also doing HIV testing? That are talking about safe sex? So it hits all three levels of what the NPP call population health.

MEMBER STOTO: So I guess from a measurement validity point of view, though, the question is, suppose somebody tests positive on a home test. Is that fact reported to a local or state health department or CDC?

CO-CHAIR JARRIS: It won't be required to be. But on the other hand, if you create an environment where people then go to clinical testing, the right thing has happened. And in theory --

MEMBER STOTO: So I think -

CO-CHAIR JARRIS: -- it could drive earlier testing.

MEMBER STOTO: So I think that actually helps the validity. So the home testing is basically an incentive for people to get into the official system, essentially.

DR. WORTLEY: I think if home testing doesn't result in people who test positive presenting for care, there's a failure there. I mean, because we want these

people to then show up for care somewhere.

And if that person who tested positive and then waited five years, and presented for care once they were systematic, that would be a problem.

I mean, they would show up as a late tester, even though in fact, they had tested earlier. But the testing would not have accomplished what it needed to.

DR. HALL: That's not really a diagnosis, home testing, right?

DR. WORTLEY: No. Not a confirmed diagnosis.

CO-CHAIR JARRIS: Bobby.

MEMBER PESTRONK: How does NQF, on the clinical side, deal with this sort of fractaling that happens in the conversation about a particular measure? I mean, now we're beginning to examine other aspects of the care system that could influence this population measure.

And yet, it seems to me from the

discussions that we've had, NQF wouldn't call for -- Or maybe I shouldn't say that. Would NQF, now recognizing that it would be useful to have other kinds of quality measures in its basket, call for these other kinds of measures in order to have a more complete understanding of how the care system's actually working?

DR. BURSTIN: That's a great question. And I like fractaling as well, of a measure. I haven't thought about it that way.

I think in some ways what we would really more so look at is whether those other considerations, like home testing, are threats to validity of this measure.

So as an analogy, we've had a whole series of discussions about the medication adherence measures that we have endorsed, in which increasingly people are using the $4 generics, and not necessarily putting those prescriptions through their health plan at all.

And so increasingly there's concern that that estimated ten to 15 percent of prescriptions, to look at adherence with certain drugs, may not be captured in health plan data anymore. Sarah knows this well.

Is that a threat to validity of any of the other underlying adherence measures? So I think for where we are right now is, I'm listening to this discussion.

The main question for me is, is the growing movement towards home HIV testing an inherent threat to the validity of this measure?

Or is it something, which I think it is given the newness of that, that that is something that we would encourage CDC to keep an eye on, you know, as the measure comes back for re-review.

They increasingly try to present data to give us a sense of what proportion of patients we might not be capturing because they're going outside the traditional testing

pathway.

CO-CHAIR JARRIS: And, Pascale, or others, can you, I mean, there must be research into the whole notion of home testing. Because there certainly was objections to that by many people, that demonstrated some benefit to home testing.

And, for example, was it specifically looked at, that home testing actually did drive into earlier clinical care? Because there's probably literature on this.

I can't believe FDA approved it, and CDC would have endorsed it if there wasn't some evidence to say home testing is a good thing.

DR. WORTLEY: And you know, I'm unfortunately not familiar with that body of literature. But I would imagine that there was information presented. And also I would guess, and I certainly can follow up and find out for you, is that I'm sure we'll be monitoring it.

The question is through which vehicle we'll be monitoring it. And we do a survey of high risk persons. And I would think that would be one way that we would monitor it. So at this --

CO-CHAIR JARRIS: So if the measure is a population health measure, basically saying, if you do well on this measure, that people are diagnosed early on. That essentially means you've created a whole environment among that population, where people seek testing and care early.

So if you were being measured on this, let's say as a kind of a care organization, and doing poorly, you would have to go out in the community to work on this issue, which is what I like about it. It really does drive people toward a population approach.

And one of those strategies you might very well specifically take on is home testing, as a way of feeding people in. So I

think it may not at all be in conflict with this measure. It may actually be supportive, and an intervention you'd want to take up to improve on the measure.

MS. JACOBSON: Elisa, this is Dawn. I'm on the call. Can I make a quick comment on this, that I would like to put on the table to discuss tomorrow as well?

MS. MUNTHALI: Sure.

MS. JACOBSON: What I liked about this measure is that it is a real population health measure, that uses a population health data set.

And I think, if you think of the table that we were doing as examples in the white paper, what you're describing are interventions that would be what we're, at least for now, calling health improvement activities, that would either be done by the clinical care system or, you know, another community based organization or a public health agency.

And so this is where the linking of the health improvement activities to the ultimate population health outcome measures that were selected would come into play. And no one's reporting like that yet.

If NQF moves in that direction, it would be the first organization that I know of that would be really getting the complimentary and synergistics. And your interventions link to the health outcome.

Even healthy people doesn't do that. They mix them up in their chapters. So I would like to put that on the table for tomorrow's discussion.

CO-CHAIR STANGE: This is Kurt. I think it's -- I also like this as a population health measure. And I think it's really an outcome measure. And the point that Bobby brought up.

I think there is a higher level of utility of a measure that lets you start thinking mechanistically about the ways that

it might be changed. I don't think to be a good measure that you have to have those mechanisms built into the measure.

And the simplicity of this, particularly for something that we don't really have a good way of getting a denominator, is nice.

I mean, if there are some ways to add some of those things on that would let you know what you would do about this outcome, that would be great.

But I don't think it needs to be a criteria. Because I think that just complexifies the measure. And it really starts to eat away at the feasibility.

CO-CHAIR JARRIS: And we don't want things complexified. So, Mike.

MEMBER STOTO: I agree with those comments. I want to come back to Helen's point about viewing this as a threat to validity, the home testing, and broaden it a little bit.

I think the way you have to think about this is that, the concept we're talking about is widespread in early screening, and widespread in early testing. Although it's measured in terms of proportion with late diagnosis. The concept really is screening and testing.

And then if you think about it that way the home testing is not a validity problem. But in fact, it's an intervention. And I think that makes it -- I'm more comfortable with it because of that.

MEMBER BIALEK: Does the validity of the measure ever become part of what gets approved as the measure? What I mean by that is, in this instance we have it looks like 80 percent or so completeness.

And if one determines that this is an important measure, and if the completeness of the reporting declines, then that would suggest there's a potential adverse population health impact. And so does the completeness

of the measure possibly become a measure as well?

MEMBER STOTO: What's the 80 percent completeness?

MEMBER BIALEK: What I was reading here is completeness of HIV and AIDS case reporting is estimated at more than 80 percent. And I didn't know if that meant --

MEMBER STOTO: So I think the question -- If 80 percent -- If there's 80 percent completeness of the original HIV test, and 80 percent of the AIDS result, the proportion late may actually be correct.

Even though there's incompleteness. So I think you have to sort of come back to the question about what's the concept that's being measured here. And it may be that that completeness is not an issue.

I need to think it through to be sure about that. But that's the question I think really is, does this completeness get in the way of this measure measuring the concept

that we want?

CO-CHAIR JARRIS: Is completeness something that is tracked on an ongoing basis? And how do you do that?

DR. HALL: So we have completeness measures for -- We evaluate our surveillance system once a year. And we have some methods that we use, like capture recapture. Or we calculate the expected numbers based on some regression analyses.

And right now HIV and AIDS reporting is mandatory in all areas. Of course there's always possibilities that completeness may lag because of personnel turnover or other issues.

And it is also really very, very high, virtually 100 percent, in areas where there's 100 percent mandated laboratory reporting, where all HIV diagnostic reports come in, and where all CD4s are reported.

As you may know, most of the AIDS classifications are these days based on CD4,

rather than on ROIs. So while there are always, you know, some challenges with completeness, completeness is very high for the surveillance system.

And the other incentive for being complete is that Ryan White funding is based on both AIDS case reporting and HIV case reporting. So it's in the interest of all areas to make sure that they have complete data.

CO-CHAIR JARRIS: What is the possibility of looking at this at less than a State level? Or perhaps you have directly funded counties or cities. But I mean, how could this be used at a health system level? A health care system level?

DR. HALL: Well the levels that can be looked at is state, city, county, census tract and by diagnostic facility. So it can be looked at, you know, at all of these levels.

DR. WORTLEY: I think that when we

think about looking at it by diagnostic facility, it's probably good to think about what kind of facilities it would make sense to look at for. For example, closed systems like Kaiser or VA. I think it's easy to see that it would make a lot of sense.

Other facilities it wouldn't make necessarily a lot of sense. Because it might just be a place where, you know, people showed up, presented once they were ill. And their testing could have happened anywhere. But used judiciously, it could be used at the health system level.

CO-CHAIR JARRIS: Matt, and then Bobby.

MEMBER STIEFEL: Perhaps this is a tomorrow topic as well. But there have been several comments that this is an outcome measure.

And it's important because we have to make a determination here about if it's an outcome measure or not in our evaluation. It

is certainly associated with an outcome.

But, you know, I might argue that it's somewhere in the causal pathway short of an outcome. That the outcomes have to do with infection rates and complications of the disease, and mortality associated with it.

And so I mean, there's an immediate question about how do we handle that. But then the broader question about families of measures on this causal pathway to outcomes.

And maybe this is a piece of a family of measures that ultimately I think we would want to look at infection rates and mortality.

MEMBER STOTO: Could I just speak to that one? I mean, Dawn mentioned that it was a population measure. And I agree with that. But I agree with Matt, that I don't think this is an outcome measure.

And I think that if we could somehow measure the proportion of people who

have HIV, who have been tested positive, which I think this is what this gets at. We would see that would be a process measure. And that would be short of an outcome measure.

CO-CHAIR JARRIS: But it's a process measure. But it's not unlike A1C, which is a proximal measure. BMI is not an outcome measure.

MEMBER STOTO: I think it's more like have you been tested for A1C, rather than what is your A1C level.

CO-CHAIR JARRIS: Well it began with, have you been tested in the last year? And then it went to six months. And then it went to the A1C level of less than seven, where all of us in clinical medicine thought we were going to be killing people, until we learned you could actually do that. And so, you know, this measure may be the same, that over time it gets refines.

MEMBER STOTO: But as you went through that progression, you switched from a

process to an outcome measure. And I think that even though this is measured in terms of disease status, in fact what it's getting at, the concept that it's getting at is screening.

DR. BURSTIN: I think we would classify this as an intermediate outcome measure. Very analogous to A1C testing versus blindness, mortality, et cetera, complications from diabetes.

Late diagnosis in general, I mean, I think that there's multiple measures that sort of look at an outcome of a late diagnosis, which I think is essentially getting at. To me it's an intermediate outcome. But it's --

MEMBER STOTO: Yes, but I don't think it's getting at late diagnosis. I think late diagnosis is the specific thing that's being calculated.

But I think the real thing that they're interested in is screening. And that's the challenge here is, that the late

diagnosis.

I mean, there's no doubt, if you look at the words it's late diagnosis. But the reason for doing it is because -- And that's the distinction I think is important here.

MEMBER QASEEM: So Michael, if it's getting to the screening, why don't we just make it the screening measure? I'm a little bit confused. Maybe I'm getting lost a little bit here.

MEMBER STOTO: Well there's two ways of doing that. One is you could say, of everybody, what proportion has been screened? And I guess you could do that with the survey questions.

MEMBER QASEEM: Well because if that's the intent, maybe that's what should be the measure. Because that's now how it's reading right now, the way it's written.

MEMBER STOTO: Yes.

CO-CHAIR JARRIS: This is more

than a screening measure. Because this is basically at the point of diagnosis. What stage are you at screening? You could be screened and then never go to care. And you've been successful at being screened.

MEMBER QASEEM: Exactly.

CO-CHAIR JARRIS: Except you're dead. So I mean, this is a little bit further downstream I think than just screening.

MEMBER STOTO: Well I agree with Matt, that this is a question for tomorrow. I don't think it's am easy call.

DR. WORTLEY: Can I just interject one thing? One of the complexities around the have you ever been tested issue, is that we're not actually recommending for the entire population of the country to be tested.

The recommendations read for routine opt out testing in health care facilities with an undiagnosed prevalence of HIV greater than .1 percent. And so we do use the NIHS, for example, to see what percent of

the population has ever been tested.

But the need for testing in Omaha isn't the same as in New York City. So we don't really know, well what should it be. And in Omaha it probably makes more sense to have targeted testing programs.

So the late tester diagnosis, what it's really looking at is the effectiveness of your testing program, however you're implementing it given where you are, which would be different in New York City and in Omaha, for example.

CO-CHAIR JARRIS: Bobby, you've had your card up.

MEMBER PESTRONK: I was going to say that it's both screening and treatment. But when I look now at the actual numerator statement, there's nothing about treatment.

Diagnosis presumes treatment, I guess. But it doesn't necessarily mean that treatment has followed diagnosis. So it's really screening then that this is looking at,

rather than treatment.

I think it's presumptuous to assume that diagnosis necessarily leads to treatment. The other thing -- So that's on that, my observation on the last comments.

Before that, I am looking at -- I mean, the outcome is certainly important. But the way at least the evaluation sheet is constructed right now, it talks about opportunity for improvement. It doesn't talk about outcome necessarily.

And there are lots of ways to think about this as an opportunity for improvement. And that's what I think is important about this measure, as a population measure.

That's because I don't know anything about HIV and screening, and rates, and what's actually going on there right now. So my question to the folks that do know is, how much variability is there right now, in where there are data available to understand

what kind of variability in this measure there might be? What's the answer to that question?

CO-CHAIR JARRIS: There is some documentation in here on that, in terms of people at high risk and low risk, and things like that. And there is quite a bit of --

MEMBER PESTRONK: Quite a bit of variability, yes. So then it's -- Yes, okay.

CO-CHAIR JARRIS: Matt, and then Elisa.

MEMBER STIEFEL: Well it's kind of a funny type of screening. It's like a Bayesian screening measure. It's screening given that the disease happens. And you don't know that when you're going to do your screening program obviously.

So the screening decision I think we're hoping to induce or create, is screening of what you know at the time, which is high risk people. And so it just seems clearly that this is a --

We're trying to improve screening

for high risk. Because you can't improve screening for known disease. Otherwise it's a cart and the horse problem.

CO-CHAIR JARRIS: But doesn't this also drive the link between screening and definitive diagnosis? It's not purely screening.

MEMBER STOTO: Well the reason for screening is to identify people who have the disease. And the sooner the better.

CO-CHAIR JARRIS: But it doesn't necessarily happen without building a system where screening leads to care.

MEMBER STOTO: I guess -- Well maybe the question is what you mean by screening. I think that it's a screening and testing, and getting a diagnosis.

CO-CHAIR JARRIS: Well I don't lump all that together in screening.

MEMBER STOTO: Yes, right.

CO-CHAIR STANGE: So this is a multi -- There are many pathways to get --

MEMBER STOTO: But not testing.

CO-CHAIR STANGE: -- to this beyond screening programs.

MEMBER STOTO: But not treatment.

CO-CHAIR STANGE: I mean, there's public awareness. There's early detection in clinical care. There's lots of ways, pathways, by which this is an overall synthetic outcome issue.

MEMBER STOTO: Yes. So Dawn is right that this really is a good example of a population health measure. And it's one as, I guess it was Pascale that was saying is that is allows different communities to approach it in different ways, depending on the nature of their population and their respecter, and so on.

And it really does measure pretty well how good you're doing at screening and diagnosing people at high risk, who are the ones who should be getting it according to the CDC recommendations, in an indirect way.

MEMBER QASEEM: But in this one they are talking about treatment. Maybe I'm -- Because they are talking about going to Stage III. So the treatment is part of it.

I don't think this is -- I'm not, again maybe -- Like Bobby, you just said that it's reading as screening. I still don't think so.

Because, yes, I completely agree with what you're saying that it's diagnosis. But then if you read the numerator it's talking about diagnosing progressing to Stage III.

MEMBER STOTO: Well but remember, the three month window is aiming at getting at the concept of someone being diagnosed with AIDS at the same time they're diagnosed with HIV.

And it's not that there's an intervention within that three month period that will stop you from getting, prevent you from transitioning to AIDS. It's that you

were diagnosed with HIV so late in your disease process that you're already at AIDS.

MEMBER QASEEM: So can I ask a question from CDC folks? Can you just help us out? What was your intent? Is it for getting to screening? Maybe I missed that part. Is it for screening? Or are you looking at diagnosis here? I mean treatment.

DR. WORTLEY: No. It's not about treatment. It's about really the effectiveness of the testing activities in a given state or community.

MS. MUNTHALI: I just wanted to remind the committee about the preliminary voting results from the workgroup. We have those up on the screen. I don't know, Paul or Kurt, are there any of the results for any of the criterion that you want to point out?

I'll also have you look at the overall recommendation and the rationale for that. And feel free to ask your colleagues for more input on that.

CO-CHAIR STANGE: I guess particularly focusing on is there anything new that hasn't come up in the discussion to focus us toward getting ready to vote on these things? So with this as yet another source of input in addition of our discussion.

MEMBER STIEFEL: This is just a question. How do we address the issue about the data presented over 12 months, but the measure is about three months? So when we evaluate the sufficiency of the evidence, what do we do there?

MEMBER STOTO: I think there's a broader question that if they could re-write their proposal, having heard this discussion, they would make a much stronger case for a lot of things. Not only that but a number of thins.

So I think that if we look at the documentation per se. And this is a much of a problem, I think of NQF as it is of CDC, it doesn't -- A lot of these points are not made

adequately. That's only one.

But I think that with everything that we -- So I would vote very differently knowing what I know, versus knowing what I saw in the documentation, is the issue.

DR. BURSTIN: Just one comment, as somebody who has reviewed way too many NQF measures over the years. I actually think this is actually quite well done.

I think the evidence is incredibly clear. They make a clear connection between early testing and the outcome. I think they provided very good evidence that the surveillance system is reliable, with a high capture rate.

I do think Matt's point is well taken. I think it would be helpful to understand what they expect the differences might be in terms of the reliability and validity of the measure versus 12 months, versus three months.

And is there any reason to expect

that those data are somehow going to suffer? But I actually see this as quite good. And actually remarkably fit the boxes well.

MEMBER STOTO: I disagree about the documentation. But I agree about the final value of the measure.

CO-CHAIR JARRIS: So this issue of three months -- And like the CDC response. Is it 12 months just to allow the period of time to clean up the data?

DR. WORTLEY: Irene, do you want to take that? Or I --

DR. HALL: Go ahead, Pascale.

DR. WORTLEY: Well the 12 months was what we -- As Irene mentioned a while ago, is the measure that has been in our surveillance report for a number of years.

And because of the desire to make it a cleaner measure, that is, a measure simply of concomitant Stage III and HIV diagnosis, that is, somebody who's being diagnosed at the time they have CDC defined

AIDS.

It was changed to this three months, which is -- The three months I think has caused some confusion. But it is meant to mean a concomitant. That's their stage of diagnosis.

And because the data are being currently prepared for release, we can't share them with you. But the vast majority of people, in terms of, you know, giving you a sense of what the data would look like. They don't really look that different.

Because the vast majority of people who were diagnosed within 12 months in fact have a concomitant diagnosis. Does that clear things up a little? We don't expect that the reliability or the validity would be any different.

CO-CHAIR JARRIS: Okay. Anything new before we go on to the voting process.

MEMBER KINSINGER: This is Linda. I just have a question. So that's actually

really helpful. But do you expect changing it to three months then to improve care? Or just to make the whole process a little cleaner, and more accurately reflecting what it is you're looking for?

DR. WORTLEY: It makes the measure cleaner. So it really -- As Irene was saying before, when we look at what's in 12 months, then the whole concept of treatment and progression comes into play.

Whereas here we're just saying, of the people who got newly diagnosed with HIV in 2012, how many of them were diagnosed so late that they in fact already had AIDS? And that's really what the measure is.

MEMBER KINSINGER: So you don't necessarily expect performance of screening to necessarily improve?

DR. WORTLEY: No, we do. Because if, you know, at the extreme example, if we were to all of a sudden screen everybody in the country, all of those undiagnosed people

out there, and we know there's 250,000 of them, would then be diagnosed. And they wouldn't be diagnosed in five years or in ten years when they might be late diagnoses.

CO-CHAIR STANGE: This is Kurt. It seems to me that the going from 12 months to three months improves the interpretability of the measure.

DR. WORTLEY: Exactly. That was the purpose of it.

MEMBER STOTO: And the only reason --

DR. WORTLEY: Irene, if I'm saying anything wrong, correct me.

DR. HALL: I'll correct.

MEMBER KINSINGER: And that seems a good outcome, to have it be clear. But I just think it's important for all of us to understand that's what we're looking for, is just simply -- Not simply perhaps. But just trying to make this a better measure, as opposed to expecting that all overall care

processes will change.

MEMBER STOTO: And the reason that the 12 month figures were presented is because the three month -- It's just a timing issue. The three month figures are not publicly available.

DR. WORTLEY: Right. If we had been doing this in six months, we wouldn't have even had to talk about the 12 month measure. We wouldn't have. All we would have talked about was the new measure.

CO-CHAIR JARRIS: So now that we've decomplexified this, let's move on to voting.

MS. MUNTHALI: Follow my own instructions. For 1A, impact, our rating scale, high, moderate, low, insufficient evidence. We have everyone in the room. I think Linda's gone. Oh, Linda, you're still there. But Sarah's gone.

MEMBER KINSINGER: Moderate.

MS. MUNTHALI: Okay. And so we

have eight high and three moderate. So we can go on to 1B, performance gap. Same rating scale. Three are missing from the -- Okay, we have everyone. Linda?

MEMBER KINSINGER: Moderate.

MS. MUNTHALI: Okay. So we have seven high and four moderate. And so 1C, evidence. This is one yes, two no, three insufficient evidence. One more. Okay, we have everyone in the room. Linda?

MEMBER KINSINGER: Yes.

MS. MUNTHALI: Okay. So that's ten yes's and one no. So we move on to the second criterion. Oh, sorry. We have to -- Sorry, I was moving too fast. Okay. So this is reliability and validity. The scientific acceptability of the measure properties.

And just to remind you, this is the extent to which the measure as specified produces consistent and credible results about the quality of care when implemented. And the measure testing must demonstrate adequate

reliability and validity in order to be recommended for endorsement.

And so for 2A, which is reliability, high, moderate, low, insufficient evidence. I think we need three more. Okay, we have everyone in the room. Linda?

MEMBER KINSINGER: Moderate.

MS. MUNTHALI: Okay. So four high and six moderate, and one low. Okay, so we'll move on to 2B, which is validity. And again, same rating scale. I think we have one missing from the room. Okay, we have everyone. Linda?

MEMBER KINSINGER: Moderate.

MS. MUNTHALI: Okay, so three high and seven moderate, and one low. Okay. So now we're moving on to usability. And this is the meaningful, understandable and usefulness for public reporting.

Rating scale, high, moderate, low, insufficient evidence, insufficient information. The timer has started. We're

missing four. We have all ten. Linda?

MEMBER KINSINGER: High.

MS. MUNTHALI: Okay. So we have six high, four moderate and one low. Okay. So feasibility. The extent to which the required data are readily available, retrievable without undo burden. And can be implemented for a performance measurement.

(Off the record comments.)

CO-CHAIR JARRIS: Let me just say that if the summary --

MS. MUNTHALI: I'm sorry, the mic is off.

CO-CHAIR JARRIS: -- of this voting feels like it would make a significant difference to split the two, you can raise that as an issue.

MS. MUNTHALI: Okay. So feasibility. High, moderate, low, insufficient information. One more. Linda?

MEMBER KINSINGER: High.

MS. MUNTHALI: Okay. So five

high, five moderate, and one low. Okay. So this is the overall suitability for endorsement. Does the measure meet NQF's criteria for endorsement. This is yes, one, and no is two. Okay, we have everyone in the room. Linda?

MEMBER KINSINGER: Yes.

MS. MUNTHALI: Okay. So --

CO-CHAIR JARRIS: Can I just raise question?

MS. MUNTHALI: Oh, sorry.

CO-CHAIR JARRIS: This may be precisely how the voting is going, but we always have one no. And I'm just -- Are all the clickers working? Just may be different people. I'm just curious.

MS. MUNTHALI: We think they're working. We're receiving it, yes.

CO-CHAIR JARRIS: Can we do a test vote? We're all asked to vote, press number one?

MS. MUNTHALI: Let me just, as a

matter of public record, just state what the vote was for this. It's ten yes's and one no. And so do you want to do a test to make sure --

MS. CHANDLER: Everyone press yes --

CO-CHAIR JARRIS: Everyone presses a one, let's make sure we get all ones.

MS. CHANDLER: If everyone -- We're going to try it one more time. If you could look to the light, you should see a green light if your remote is working. If anyone sees a red light, please let us know.

(Off microphone comments.)

CO-CHAIR JARRIS: So 2020. Well anyway that is the first time we've actually endorsed a measure. So this is a historic moment, right? Well first time in this process, second stage. So let's go on to -- That's right, 2020, adult current smoking prevalence. And who do we have? Dawn Jacobson, from Legacy. Are you on the line?

MEMBER PESTRONK: Excuse me. In the SharePoint there are two 2020s. The one we're looking at is the updated one. Is that correct?

MS. CHANDLER: Yes. That's correct.

MEMBER PESTRONK: Okay. Thank you.

CO-CHAIR JARRIS: Donna, are you on the line and on muted?

MS. MUNTHALI: Anika?

OPERATOR: The line has disconnected.

MS. MUNTHALI: Oh, her line disconnected. Okay. So I guess we can start talking about the measure. And if Donna wants to join later when we have questions, we'll try and get in touch with her.

CO-CHAIR JARRIS: Okay. So this is Amir?

MS. MUNTHALI: This is Amir and Kurt.

CO-CHAIR JARRIS: Kurt.

MEMBER QASEEM: Okay. So I'm not sure if I have the updated one. Maybe I didn't download the updated one. Is that the one where fix on the numerator and denominator show? It probably did. Okay.

MS. CHANDLER: Yes. So just in the last few days, since may of you inputted your results, we received an updated form, which is on SharePoint. And I'll circulate that right now if you don't have it.

MEMBER QASEEM: So this measure is about the percentage of adults who are currently smoking. And it's a pretty simple measure actually, I think.

And I do think it's a very useful measure, in terms of to find out what's going on in your community to figure out of course, the resources as well as the plans you're going to put into place to address this issue.

So in terms of the evidence that was presented, you all know that smoking can

be beaten. There is enough there regarding this.

My only thing, that I think I raised in the group as well. Isn't this information already available? And that may be just me. And maybe in the area where I'm in, where I've worked. I always felt like this information was already there.

So I wasn't really sure what we were adding. Or maybe, there's probably a lot of communities that are out there that this information is not available. So it's going to be extremely beneficial in that case. But that's my only thing. Otherwise it's a pretty good measure.

CO-CHAIR STANGE: Sure. I think it's a great population health measure. It's feasibly collective. It's not -- You're not able to get granular geographically on it.

And of course there's other kinds of tobacco that are an issue. But it's already being collected. It's useful to

monitor.

We have data going back decades. So it's useful to look at trends. So I guess the only issues would be harmonization with other measures. But I didn't see a problem with it at all.

CO-CHAIR JARRIS: Madeline.

MEMBER NAEGLE: I'm wondering if the maker of the measure would speak to the reasons for the exclusion of people in the military and for people who are institutionalized?

Since the rate of smoking among people with mental illness is the highest. People in substance abuse and psychiatric facilities have the highest.

So to use a measure which excludes two groups. Military I would also submit are at risk for addiction. So would the maker speak to that please?

CO-CHAIR JARRIS: I don't know that they're on the line. We're calling them.

So maybe we should hold.

MEMBER STOTO: I think I know the answer to that one.

MEMBER NAEGLE: Okay.

MEMBER STOTO: I think it's because they're not in the National Health Interview Survey, which is what they're talking about here. So I think --

I mean, I agree that this is an important thing to measure. There's no doubt about it. But the documentation is really about its implementation in the National Health Interview Survey, rather than about it as a measure.

So, yes. I mean, what about the Behavioral Risk Factor Surveillance System, and so on? And lots of other surveys that ask about smoking?

I think that we should be endorsing the measure based on the definition of smoking to be used in surveys, rather than the particular thing.

So there's lots of things like, you know, what does it count to be a smoker? And that's not addressed here. I think typically you have to be -- I think there's a standard definition of that, that's used in the surveys. I forget what it is.

CO-CHAIR JARRIS: Smoked one or more cigarettes --

MEMBER NAEGLE: More than ten cigarettes a day.

MEMBER STOTO: Something, well I don't think it's that much. But there is a standard.

CO-CHAIR JARRIS: Jason.

MEMBER SPANGLER: I had a question. And this is directed mainly to NQF. There are several other tobacco measures. And some of them seem like they should be with us. And they're not. And I'm wondering why.

And not necessarily phase -- I don't think it was brought up in Phase I. Because I assumed they'd be talked about in

Phase II. But there aren't any in Phase II.

And I know that some of them haven't passed through our steering committee, but are in the behavioral health steering committee. Just wondering why, if we know why that was.

And because for a lot of us in prevention in public health, they are population health measures. And whether there needs to be some harmonization of those with this measure. And how that would work when there's two different committees working on different types of measures.

DR. BURSTIN: You're absolutely right. So these are measures that could have fit in either place. Because we knew there was going to be a whole large set of measures that were hospital based, we thought it might be best to keep them with the behavioral health measures, including mental health, substance abuse, et cetera.

That committee -- Actually

Madeline happens to be serving on both committees. So she is your bridge to both of them. Thank you, Madeline. And those measures were just reviewed. And I can tell you that there was a --

And actually maybe one useful thing for us to do before we conclude is, we'll bring you the, at least the description and the specifications of the measures preliminarily approved. Actually going out to comment.

I think it just began yesterday. So there is a hospital based screening measure, that the Joint Commission put forward, that narrowly was approved. There is a clinician level measure, where clinicians --

It's a question. Did you ask the patient do they smoke and offer assistance with quitting as a combination measure. And then there's a third measure.

And actually NCQA's here. Dawn may want to speak to this, which is actually

a series of supplements to the clinician group health plan, CAHPS? Thank you. The health plan CAHPS tool, in which there are a couple of specific, smoking specific survey measures added there.

So the issue is, none of those get to the population level. I still think it's reasonable for us to consider whether the definitions are a lie. But it's a little different because one of them is a provider answering the question.

One of them is a survey though, at a health plan level. And one of them is a new institutionalized patient question. So I think the harmonization issue is definitely worth exploring.

CO-CHAIR JARRIS: If I could just come out, just in terms of the strategy? If we're looking at this process, we'd like to have a smoking measure, as opposed to not having a smoking measure.

And one of the issues we've had is

that people didn't come forward with measures. So I would suggest we adopt what we get. And then in the next process we'll be looking for a best measure among a group of measures.

But rather than hold off and have no measure, let's pick this one and hope that will tweak the other people who didn't submit their measure.

CO-CHAIR STANGE: If I could just follow up on that? To the question, what does harmonization mean in this context? Harmonizing this with the clinical measures doesn't make sense.

We're talking about harmonizing with other population measures. And we don't have those on the table. So what does -- I mean, do we just say we need to -- This is fine, but insufficient for sub-populations. And there may be other complimentary or better ones.

DR. BURSTIN: This is the way we try to explain harmonization versus competing

measures. I don't think these measures are competing. They are a different population, different denominator, different patient population. I think they are potentially issues of harmonization, because it's the same measure focus.

So in fact, if this measure at the population level asks about hundred cigarettes. And the other measures don't. I guess to me there's an opportunity there for harmonization. I agree it's not clear it has to happen right now. But I do think it's something worth exploring.

CO-CHAIR STANGE: Will the NQF be accepting other population measures? Because it would be -- To Paul's point it would be fine to have this out there.

Then people that know of other measures could say, wait a second. They get sub-populations that are excluded for logistical reasons, or whatever. Would there be a process for people to submit other

measures?

DR. BURSTIN: Yes.

CO-CHAIR STANGE: Okay.

MEMBER KINSINGER: And this is Linda. Just maybe again for the discussion tomorrow. But in thinking about the forms. I found the evidence presented by the developer here about why smoking is bad for you to be not very useful. Nobody's arguing whether smoking is bad.

I think the evidence that we're asking for from the developer is how well can this be measured? What's the impact of using this measurement? How does it lead to change?

So if there could be some clarification about, you know, what sorts of evidence we're looking for, it seems to me that would help.

CO-CHAIR JARRIS: Bobby?

MEMBER PESTRONK: A couple of things, just to go back to what Mike said. In Sections 2A 1.3.7 and .9, it gets at the

exclusions. That's the description that they simply aren't collecting those data. So they're citing them as an exclusion for this particular measure.

But this is then looking through the other end of the tube, if you will. Because those populations have the highest rates of tobacco use, some link to a stratification, which provides the data for these other populations, is extraordinarily important in understanding the prevalence of smoking in a population.

Because without those, you lose track of these populations, which actually should be the target of interventions. But because the measure that's being used excludes them, gives you a false picture of what's going on in the entire population.

CO-CHAIR JARRIS: Couldn't a complimentary measure handle that?

MEMBER PESTRONK: Yes.

CO-CHAIR JARRIS: Because see the

number of people with mental illness who are institutionalized is minuscule.

MEMBER PESTRONK: Yes.

CO-CHAIR JARRIS: And it's a temporary population in most cases.

MEMBER PESTRONK: I would vote to endorse this measure. But I think it's important in the context of the conversation that we had before to look at this through the other end of the tunnel.

And say to ourselves, you know, where are those data? Where is that measure about these two populations? Because that's, it is a minority population in the way that you describe it. We should know what's going on to reduce smoking within those smaller population groups. So that's just a comment.

The other question I had is in Section 5B, where I was looking at the competing, the description of the competing measures. And I just wondered why the point of contact for Legacy was Tom Friedan?

MS. MUNTHALI: So Legacy is the developer. And the CDC is the steward of the measure. I don't think --

CO-CHAIR JARRIS: I think Matt, I think and then Mike.

MEMBER STIEFEL: The notion of looking at total population only is kind of a slippery slope. Because it would exclude a lot of the other measures proposed for this group to review.

The NCQA measures that are health plan performance measures. So it's a pretty important topic. And, you know, using the language of the white paper, there's this notion of a geopolitical population with sub-populations within it.

I don't think we're looking at only geopolitical populations for all of the measures here. And harmonization just becomes that much more important, where there may be a number of sub-population measures for smoking. And in fact --

So the broader question is, I think a lot of the measures we're looking at are about the boundary or intersection between public health and clinical care.

And how we may roll up clinical care measures for a geographic area, to begin to look like total population measures. So maybe this is just a tomorrow question. But it's an extremely important topic.

MEMBER STOTO: I was going to say the same thing. And I would just add one more thing. To the extent that, you know, we're moving towards accountable care organizations, and so on, I mean, that just forces us to think about this intersection even more.

And it's an absolute essential that measures we talk about for the whole population, geographical populations, are consistent with the ones that we use for health service organizations.

MEMBER STIEFEL: And it's really important. Because without getting this

information from clinical care, we rely on things like these very high level, population level surveys that are getting worse and worse over time, from BRFSS to National Health Interview Survey. And so we really need to get this information from the clinical system.

CO-CHAIR JARRIS: I think they're all complimentary. And the population's relative. The population may be me as an individual, my family. If you're a clinician it's what is my patient panel.

If you're a insured it's who is my enrollees. If you're a county it's who lives in my jurisdiction. I think that all those things are complimentary based on sort of your line of vision.

MEMBER STOTO: I think that's absolutely true. I was just reacting, maybe Matt too, to Helen's point that we're talking about population today. And these other definitions of population are in some other part of the discussion.

DR. BURSTIN: The question I guess -- I didn't realize this was only at the national level. This doesn't, there's no capacity to bring this down at all to lower levels of aggregation?

MEMBER STOTO: Well that's why I think that we need to think about the question, as opposed to the particular implementation on NIHS. On NIHS it is only at the national level. But the Behavioral Risk Factors Surveillance System is at the state level.

And New York City has its own survey. And California has its California Health Survey that can do it in a lot of geographical areas. So there's lots of different ways that this can be implemented.

DR. BURSTIN: But do all those surveys use the same questions?

MEMBER STOTO: Well I think that they tend to.

FEMALE PARTICIPANT: And the

definitions?

MEMBER STOTO: Well probably not exclusively. I think that there has been a lot of attention over time to try to get them to formalize those things. I'm sure it's not perfect.

CO-CHAIR JARRIS: Sarah, did you want to say something on that.

MEMBER SAMPSEL: I would -- Just to answer Helen's question. They don't use the same questions specifically. So you can't really compare the results from BRFSS at the county or state level to the national NIHS results.

There is some capability of drilling down on NIHS results. But they're rarely published. And still you can't get down to the county level like you can the BRFSS.

CO-CHAIR JARRIS: Ron, I think. And then Amir after that.

MEMBER BIALEK: I have a comment

and a question. First of all, did anybody from Legacy end up on the line? Okay. Okay. Just the question that was raised earlier, you know, why this particular measure?

There was some very practical considerations was given the time frame. This was probably the easiest measure that Legacy could develop and provide the documentation within the time frame that they had. They are interested in other measures.

And I don't know if it's appropriate for the committee to provide some guidance, if you will. But that's something that I think they would look forward to.

There was some discussion about taxation measures, and discussion about indoor air types of measures as well. So that's the comment.

The question is really for NQF, which is, when we endorse a measure, are we endorsing the measure and the source? Or are we endorsing the measure where there could be

multiple sources?

DR. BURSTIN: It's actually very dependent on the way the measure is submitted. And the reliability and the validity data they submitted.

We've had this discussion very similarly with some fo the child health surveys that have been submitted to us and endorsed.

And those are conducted in a way where they have done, for some of the items, a good deal of research showing that those data can be used with those survey items, a group of survey items we pulled out from the larger child health survey to be used at lower levels of aggregation.

So it was more so that they pulled out a series of measures out of the larger child health survey. And at the state level those data are dependent currently on that child health survey being administered.

And it was a major point of

discussion at the Board of what is the utility you need to wait on the national source of data to provide, you know, information. And that only went to the state level.

I think this one may be a bit more controversial. Because I think the question, particularly when you get to usability is, what is the usability to drive quality improvement at a --

Or what's the accountability for the nation? Short of Regina. I mean, it's just this is a tough one I think in some ways, if it's only at that level of analysis.

MEMBER STOTO: But I think that Legacy probably didn't understand this discussion we were having. And if they had they probably would have said, here's a definition --

Here's a smoking question. It's been used on NHIS. It's been used on BRFSS. It's been used here and there. I mean, lots of things our questions are asked differently

on those different surveys.

But I think this one actually is asked the same in all those surveys. I'm not totally sure about that. But I bet that they could make a case for this question as implemented in population surveys. If they had been asked to make that case.

CO-CHAIR JARRIS: Amir, did you still have a comment or question?

MEMBER QASEEM: Sure. I was just going to ask about the age. It says 18 and over. And I understand it's because of the data sources and all of that.

But shouldn't we think about younger age? Like when we were talking about HIV it was 13 on. Smoking starts pretty early. And also most of the intervention's varies.

And I think that data is not available and would be very helpful. The earlier you have interventions it seems to work better.

MEMBER KINSINGER: Can I just respond to him? I think that's part of the same thing that we were just talking about. It's that because NHIS -- Actually NHIS is the whole population. But BRFSS is 18 and over, if I understand it.

MEMBER QASEEM: NHIS is 18 and over too. But they ask questions in there about the children in the household. But they're not being a force.

MEMBER STOTO: Okay. So it really has to do with the specifics of the implementation, as opposed to the concept that's being measured here.

CO-CHAIR JARRIS: So this is Donald Rumsfeld's comfort. You go the CSAC with the measures you have and not the measures you wish you had. Is that it? So Bobby did you?

MEMBER PESTRONK: The 2A13 says it's conducted among the adult 18 and older. So that may be why the specify it that way.

Helen, I wanted to ask you to just say again what you just said, when it gets to the Board level.

I mean, this then -- This is a measure that has applicability at the national level. But doesn't necessarily have applicable at the state or the municipality or county level. And so is that seen as a weakness of the measure?

DR. BURSTIN: I think initially it was. And I think we've sort of gone past that. We've had a lot of discussion as part of this committee, shared with the CSAC as well, in terms of that there are accountable entities in public health. Accountable entities at various levels, state, county, et cetera, community.

And I guess the question would be, at a national level, does that hold? And I just think it's a question for us we need to really think about. But I think will get up, I know will get brought up as this measure

moves forward.

CO-CHAIR JARRIS: It's also interesting timing. And I would agree, at the national level I mean, it would be nice if Congress cared about this. But they know this data close enough at the national.

And it hasn't led to any improvement at that level. But the BRFSS is about to go through, or has gone through a revision to incorporate different ways of weighting the populations, and to incorporate cell phone surveys.

So we're just about to have announced new BRFSS data, which will actually show higher rates of tobacco than we've ever seen before. Because you're now sampling population that are lower income, and things like that with cell phones, not land lines.

So it's going to be a very confusing time to have a smoking measure. And that probably is being released imminently, isn't it? So boy this sure wouldn't be the

one I'd pick to endorse first. But if it's all we got, what do we do?

MEMBER PESTRONK: Does somebody have to claim accountability from the Board's perspective? I mean, one might argue that the expenditures that CDC is making right now for the web and television commercials on smoking are an implicit claim of accountability. I mean, is that what --

DR. BURSTIN: I don't know. I just think this will be -- We've never had a measure that came in only at a national level.

MEMBER PESTRONK: Yes.

DR. BURSTIN: So I think it's going to raise questions, particularly as you get to the usability criterion here of what is the usability for accountability and quality improvement? Just a question.

MEMBER PESTRONK: You know, this is a little bit of a tease. But it's kind of good that this is suggesting that there ought to be accountability at the national level.

Because there's certainly a push to require accountability at the state and local level.

CO-CHAIR JARRIS: This one of the --

MEMBER PESTRONK: It's part of the answer for our question.

CO-CHAIR JARRIS: This issue of accountability is one of the refrains that comes up constantly about there is no accountability --

MS. JACOBSON: Helen, this is Dawn. Can I make a comment on my previous practice on that in L.A. County?

DR. BURSTIN: Why don't we let you guys call it when --

CO-CHAIR JARRIS: Go ahead, please.

MS. JACOBSON: Or Kurt or Paul. If this is to be meaningful at the local level, where I think change happens, then this is not a helpful data source. We never even referred to it in L.A. For adolescents we

used the Youth Risk Behavioral Surveillance System.

And I think the reason the county health rankings is at least marginally successful is because it uses BRFSS data, which gives meaningful data for state and local practitioners to make influence change. So I'm really concerned about using NHIS. I don't think it's relevant in practice.

CO-CHAIR STANGE: I think one of the important things that we did the last time we were together is try to expand the view of what accountability means. And to expand the idea to include multi stakeholder groups.

I think because this is at a national level it actually, as Bobby said, raises the gaze to that level. But I agree with you that it doesn't have -- That's actually a good thing.

But I think we want to keep that idea going forward that we don't want to limit ourselves to a single organization or entity

that would be accountable.

The transformative thing that population health measures could do would be to stimulate multi stakeholder groups at various levels that are actionable. And there's limit to that to the national level.

MEMBER PESTRONK: Again, how do we go back to ask for measures that would use the ones that Dawn has just suggested as sources of data. We have to, I mean, I think we should ask for them. But that's not part of the typical process.

DR. BURSTIN: We frequently go back to developers for additional questions. And I guess if the committee feels like it would be more appropriate to bring a measure down to lower levels of aggregation to be more usable, we could --

I mean, I'm not sure if the folks from Legacy have joined us yet. But they've been working with CDC. CDC is the source of both of those surveys. I was just looking up

the items on BRFSS. They're not the same, but they're not dissimilar. It's just the question is --

MEMBER PESTRONK: It's not more. It's a both and, rather than either or, is my point. That this is good on its own. But to address the concerns that Dawn has raised, using the other data sources make it useful for other folks.

MEMBER STOTO: I think that the fact that this is at the national level is a bit of a red herring. I think that if Legacy Foundation understood the rules better, they would have said, oh let's put forward a survey question that can be used in BRFSS and other things that can show up at the county level and so on.

And I think so the bigger question is that we don't, we haven't yet communicated. Maybe because we haven't yet decided what we really want.

And this question about

accountability I agree is central. And I don't think that really is very well reflected in any of the documentation that we have yet.

CO-CHAIR JARRIS: So this notion of a measure, does it apply to -- And it will be extremely helpful to have somebody, and I think this has been said, put forward a way of measuring smoking that was then endorsed. And then could be used in all these various different surveys.

Because right now then it's just noise. Because everyone defines it slightly different. I mean, does NQF do that? Do they say, okay here's your question, no matter what survey you're using? That would be helpful to start harmonizing all this disparate information.

DR. BURSTIN: Right. When it's been validated that there is a, you know, a gold standard of this is the way to do it. You know, again, we had a similar issue of medication adherence.

There are multiple ways to measure adherence. We had a whole exercise several years ago to say, based on the available evidence this is the preferred method.

I just, the question is, given the way it's asked so differently in a clinical setting versus a public health setting, you know, is it still logical to ask in exactly the same way? Maybe we can grab the actual survey items used in the CAHPS tool that NCQA --

CO-CHAIR JARRIS: I think to harmonize it in the population health measures would be helpful. And the way you ask it is a doctor to your patient sitting in front of you could be very different.

DR. BURSTIN: I agree, yes.

CO-CHAIR JARRIS: I guess we had Matt.

MEMBER STIEFEL: It's basically what you said. But I think it raised a question about the role of this group, or the

role of groups like this.

And is the approval process always a measure and specifications, and an instrument? Or can we decouple those and talk about a measure and its specifications that could apply to --

And that would be I think an extremely important contribution. And potentially even to be able to bridge the gap between clinical measures and population health measures.

CO-CHAIR JARRIS: Probably as you think about it, the definition I've seen is, have you smoked one or more cigarette in the past year? Because that's not how I ask my patients if they smoke.

I'd say, are you a smoker? And who the heck knows what that means? So it would be interesting to standardize it in both cases. Bobby is your, is that a laggard flag?

MEMBER PESTRONK: The information that's provided in this form that we're

looking at, the question is, have you smoked at least 100 cigarettes in your entire life, from NHIS at least.

CO-CHAIR JARRIS: I've smoked 99. So I guess I'm off the hook. Okay. We don't -- Any questions on the phone? And I assume we don't have Legacy or someone from CDC here.

OPERATOR: At this time in order to ask a question, press star then the number one on your telephone keypad. At this time -- Excuse me, you do have a question from Mary Pittman.

MS. PITTMAN: It's not so much a question as I've been listening to the conversation. I agree. And I can't always pick out who's voice it is.

That having a measure at the national level, even though we may not be able to harmonize immediately all of the sources of data from the local level. It's really important to lift up that there needs to be a national accountability.

And I think, Helen, you're saying that that's a little different from the usual process. But if you could start with that and then push down to get the harmonization of the way the questions are asked, is that feasible?

DR. BURSTIN: I think there are just a lot of unknowns here. I don't think we know. I think this is why we're doing this project. We'll put it out for public comment. We'll figure out what the various forces our there say.

I will say having pulled up BRFSS that I think BRFSS is a lot closer to the way we ask these questions in practice, which is, you know, are you a current smoker? Do you smoke every day, some days, former smoker, never smoked?

So you know, it may just be a question of whether, in fact, the BRFSS survey may be a better fit as something that could be harmonized moving forward, rather than NHIS.

CO-CHAIR JARRIS: Mary, we'll move

you over to the open line.

MS. PITTMAN: Thanks.

OPERATOR: At this time there are no further questions.

MEMBER STOTO: I would like to propose that we actually postpone voting on this one, and ask for clarification. I mean, it strikes me that what we should be asking for is a survey question that can be implemented in surveys at a lot of different, for a lot of different populations, including healthcare service population.

CO-CHAIR JARRIS: Then I would add to harmonize at the different, at least geographic levels, if not down to the health system levels.

MEMBER STOTO: I just don't think that the people who put together the documentation understood this. And we probably didn't even understand it ourselves. I don't want to be critical of them.

CO-CHAIR STANGE: I wonder if, I

don't know if we're ever going to have a more informed vote than we have right now, while this is in our heads, and we're face to face.

And if we're trying to get a toe hold to move thing forward -- I would wonder, Mike, if it would better to go ahead and approve this, with the idea that we don't even have it on the table in front of us, the types of measures that it needs to be harmonized --

MEMBER STOTO: I wouldn't be able to vote yes for it.

CO-CHAIR JARRIS: What's the process? So if we go back for clarification, and they come back with something, will we even be here? I mean who will respond to that?

DR. BURSTIN: We'll just do it off line. We'll bring it to you. You'll have a conference call. We'll have you do Survey Monkey to vote. It's fine.

We routinely go back to developers, have them tweak, bring back,

respond, have the committee weigh in again.

I think the other question is, do we actually need to actually touch base with CDC is my question, since this is based on the NHIS.

MEMBER STOTO: And also, CDC is not just CDC --

DR. BURSTIN: I understand. It's --

MEMBER STOTO: BRFSS --

DR. BURSTIN: Right.

MEMBER STOTO: -- and NHIS are two very different parts of CDC.

DR. BURSTIN: I get it.

CO-CHAIR JARRIS: What about other -- I mean, what happens in terms of additional measures coming in? Is this call closed, and then we periodically there will be another open call for measures? Or can things just come in at any time and --

DR. BURSTIN: Again, some of that is evolving. I'm not sure exactly what our

time line is here. I don't know how quickly CDC could get something into us that's sort of more, you know -- Or even Legacy working with CDC to potentially even modify this one. Or modify the certs. We'd be happy to look into that.

The other thing is, I mean, that we will expect to be having, you know, many projects over time looking at these areas. So we can easily bring them in. Our thinking is usually once we kind of go past -- And I think a lot more question to be answered tomorrow.

A new area like population health, like we've done resource use, the expectation would be in any projects going forward we would welcome population level measures as well as clinical.

CO-CHAIR JARRIS: I think that's something that we really need to talk about tomorrow.

DR. BURSTIN: Yes. I agree.

CO-CHAIR JARRIS: Because there will be a time at which we can birth this thing and give it to every committee. But I think it's probably premature.

And one thing to consider would be, knowing what we just learned from this, do we actually then want to put out subsequently another call that is a more informed call?

MEMBER NAEGLE: I just had a question for Helen. So thinking about how we'd want to harmonize this with populations, specifically sub-populations.

Is that going to be forthcoming anytime soon as we move that agenda? So that at some point we could look at this in a more comprehensive way.

DR. BURSTIN: Yes. I mean the Behavioral Health Committee looked at the various smoking measures. This was not part of it because it had not been evaluated yet.

So we could certainly provide to you that comparative table of the different

measures we already have, and the way that's asked, and share that with you.

And I think the timing is good. I just don't -- Yes, I think the question is, I'm not sure they have to be exactly identical for now, given where we are as a country.

There's just a lot to do here. And I think there are going to be clinical measures as Paul pointed out. They're going to be asked differently.

But I guess the question is, if the health plan measures get aggregated up, are we going to wind up with just very different numbers if they aggregate up to states, for example. I'm just not sure.

CO-CHAIR JARRIS: Ron, then Bobby.

MEMBER BIALEK: I'm not real clear on what we would be asking Legacy to do. is it two things? Would we be asking them to suggest a particular question? And would we also be asking them to suggest a particular data source for prevalence?

Because I think we want to have a measure of prevalence that will be adopted, or endorsed. So is that the two things that we want to do? or do we want to do one, or not?

CO-CHAIR JARRIS: Yes. I sort of see that one is, what is the question? And can that be standardized? And then the second is the means to do that at sort of a national enterprise level. Federal, state, local, and then perhaps even health system. Bobby, did you have a --

MEMBER STOTO: So it's not a single survey, but it's something that a way of implementing it is --

MEMBER PESTRONK: I'm more troubled by the actual question that's being asked in the National Health Interview Survey. And whether that's actually generating what the measure says it's generating.

Because the description of the measure that's proposed is the U.S. population that currently smokes. The questions that are

asked to supply the data are whether you've smoked 100 cigarettes in your life.

That's not producing whether you currently smoke. And actually, if that -- you could never become a non-smoker if that's the question being asked.

MEMBER STOTO: That's the definition of ever smoker, rather than current smoker.

MEMBER PESTRONK: Yes.

MEMBER STOTO: So they didn't even pull the right question.

MEMBER PESTRONK: Right.

MEMBER STOTO: I mean, the right question is on NHIS, but they seemed to have pulled the wrong one.

MEMBER SAMPSEL: Actually I need to find it again. But somewhere in this documentation it says that the actual numerator is do you currently smoke now, out of those that say they've ever smoked 100. And it's somewhere. I just have to find it

again.

MEMBER STOTO: But that's not right either. That's the proportion of ever smokers who are current smokers. What we really want is the one --

MEMBER SAMPSEL: No. It's saying now. So if you go to --

MEMBER STOTO: That's the right numerator, but that's wrong denominator is what I'm saying.

MEMBER SAMPSEL: So if you go to 2A 1.2, the time period for current tobacco use is defined by survey respondents who endorse that they now smoke cigarettes every day, or some days.

MEMBER STOTO: That's right. But the denominator should not be the ever smokers. The denominator should be the total population.

DR. WORTLEY: That's in 2A 1.4. It says the denominator is the adult age 18 or older population of the U.S.

MEMBER SAMPSEL: Right.

DR. BURSTIN: According to the NHIS website the first question is a screener. Have you ever smoked at least 100 cigarettes in you entire life?

There is a follow up question about current smoking practices, which then asks do you smoke now? And then, I'm sorry, they have then transitioned later -- I'm sorry, it used to be that. Now it says, do you smoke every day, some days, or not at all.

So there may be more consistency to BRFSS than we know. This just doesn't describe it very well. So I think we really need to get a handle on how NHIS and BRFSS relate here. And are the questions similar or dissimilar?

And I'd also like to pull the questions that are being used as part of the health plan CAHPS supplement. Oh, do you? Yes, that would be great. Thanks, Dawn. That would be helpful.

MEMBER PESTRONK: You know, in the documentation here they may just have inadvertently pulled the wrong question. Because what they're showing as the questions asked in NHIS is two identical questions. Right. Cut and paste problem.

MS. ALAYON: Would it be helpful for us to just -- I'm Dawn Alayon. I'm from NCQA. I am the measure champion for medical assistance for smoking and tobacco use cessation

Would it be helpful if I describe the measure first, before I get into the questions, so you understand what the measure's about? Excuse me? Just briefly.

This is definitely part of the CAHPS survey and the CAHPS survey is part of AHRQ. And we have three different indicators advising smokers and tobacco users to quit. Discussing cessation medications and discussing cessation strategies.

So for each indicator we have a

set of questions for each. And the questions are part of the -- So we have a set for commercial and Medicaid. And a separate question set for Medicare. So when we get to the numerator we can get to the questions.

So for example, for advising smokers and tobacco users to quit we have a question of, in the last 12 months how often were you advised to quit smoking or using tobacco by a doctor or other health provider in your plan.

For discussion smoking cessation medications we have a question of, in the last 12 months how often was medication recommended or discussed by a doctor or health provider to assist you with quitting smoking or using tobacco. Examples of medication are nicotine gum, patch, nasal spray, inhaler, or prescription medications.

And for discussing cessation strategies, in the last 12 months how often did your doctor or health provider discuss or

provide methods and strategies, other than medication, to assist you with quitting smoking or using tobacco. Example and methods that apply are telephone helping, individual group counseling or cessation program.

DR. BURSTIN: Is there a screener question?

CO-CHAIR JARRIS: Yes. Those are all clinical questions.

MS. ALAYON: A screener question?

CO-CHAIR JARRIS: They aren't population questions --

MS. ALAYON: No, not at all.

CO-CHAIR JARRIS: -- nor are they assessing the prevalence of smoking --

MS. ALAYON: Yes.

CO-CHAIR JARRIS: -- along whatever the relevant population --

MS. ALAYON: Right. This is a self report. This is not at the population - Well it's not population in the sense of the CDC measure. So it's just very specific to

the patients.

MEMBER STOTO: But those questions only make sense to ask a smoker --

MS. ALAYON: Smokers, right.

MEMBER STOTO: So the previous smoker, do you smoke?

MS. ALAYON: Do you smoke? Well that's --

CO-CHAIR JARRIS: For example -- Yes, you're right. The population question was, do you smoke?

MS. ALAYON: Right.

MEMBER PESTRONK: But then in the context of the plan, it's did you ask the patient if they smoke?

CO-CHAIR JARRIS: Among representative sample of the plans members.

MEMBER STOTO: No. This is the survey of the patients. So not of the physicians, this CAHPS survey.

MS. ALAYON: Yes. This is patient's self report.

CO-CHAIR JARRIS: Okay. So let's take stock of where we are and decide what our next step is. I believe what we've talked about is gong back to Legacy/CDC to ask them if there is a measure that they can propose that could be consistently applied no matter what the measurement tool was, or the means of doing the questioning.

And then secondly a measure that could be harmonized at various populations. National levels, federal, state, local levels. And then perhaps the third question is, and could that in fact be utilized? The same question at a clinical level, as at a population level. Does that capture?

MEMBER PESTRONK: Except that I would be comfortable if they came back and said here's our proposal for each as well, federal, state, or local. For the reasons that we discussed here, which is it makes sense to have a national measure.

And until the questions can be

harmonized there's no reason not to have a national measure which draws from a source which gives valid, which has valid and reliable data to support. And the same would be true at the state level and at the local level.

DR. BURSTIN: I mean, it certainly seems logical that any survey items should be harmonized. I did find it. I'm sorry, just to interrupt briefly.

The first question sounds a lot like BRFSS's, which is do you now smoke cigarettes or use tobacco every day, some days or not at all. That's BRFSS. So I think that's, my guess is that's exactly how that got into the CAHPS tool. So we can confirm that.

CO-CHAIR JARRIS: Bobby, it may be that intermediate steps come back with separate levels of the questions. But at some point it would be real nice to know that if you aggregated all your counties that was your

state. And if you aggregated all your state, that was your national number. And that they weren't all variations.

MEMBER PESTRONK: There's no question in my mind that that would be better. But until we're there, there's no reason not to have the other, right?

CO-CHAIR JARRIS: Okay. Mike.

MEMBER STOTO: I'm going back a long time ago when I worked in this. And my recollection back then was the office of smoking and health, which I think is still part of CDC, had made a big effort to try to harmonize the asking of these questions. I don't know whether they're still active and doing these kind of things or not. But they would be a good contact.

CO-CHAIR JARRIS: Yes. When we say CDC we better ask the right part of CDC.

CO-CHAIR STANGE: I agree with Bobby that asking to do multiple measures would be ideal. We don't really have a place

on the form, I think, for putting the measure into context.

And so if there's some way of informally communicating to say that they get bonus points if they at least put whatever measure or measures they propose into the context of other population health measures that are usable at different levels, that would be really helpful to the, whoever is looking at this in the future.

CO-CHAIR JARRIS: Okay. Have we done what we need to on this measure? And then in that case, we have an opportunity to go for public comment for a few minutes. And any members of the public --

CO-CHAIR STANGE: Even before that, I just want to let people, both in the room and outside the room, what I've put up here. We have, for the discussion tomorrow we have some good general questions.

There's a lot of -- People have used the phrase, oh that's maybe a discussion

topic for tomorrow. That's come up a lot. And so to compliment the more general questions tomorrow, if we have some of these specific questions that come up in the course of doing this kind of work I think that would be very helpful.

So I've put up what I've heard as a list here. And I'm going to stick it on the wall because I've filled up this sheet. But I think, what I'd like to suggest as a process is, any time in the course of the discussion, or just a thought that pops into your head, if you have things you'd like to get on the table for the discussion tomorrow, let's capture those in real time to inform our discussion.

And I think maybe at the end of the day, both for those from the committee on the phone, and for public comment, we could have a moment just for people to suggest things for the larger discussion tomorrow. I don't know if there's any comment on that.

CO-CHAIR JARRIS: Mike.

MEMBER STOTO: I think it's a good idea to do. And I'd like to add something to it if I could?

CO-CHAIR JARRIS: Sure.

CO-CHAIR STANGE: Anybody can add anything at any time, yes.

MEMBER STOTO: Okay.

MEMBER NAEGLE: I was just going to comment that I think this is a great example of how we as a committee can make recommendations or not act on things with a goal of improving the quality of reporting. This is such an important issue for us.

And just thinking of the perspective of my work in addictions, how we have functioned really without sufficient and accurate information about this particular problem and many other problems that we're going to be looking at in relation to prevention and changing lifestyle behaviors.

So I really would like us maybe to consider there are times to maybe go for more

quality. To say, yes, maybe we need to do the optimal. And maybe it isn't just sufficient to say let's use what we have until we get there.

Because if we have the power as a committee to say, come back with something better before we go forward, maybe we should use that.

CO-CHAIR JARRIS: Which raises for me the issue of -- And some of these issues in smoking is a clear one. Do we also want to ask them for, when there's area of clear disparities for a measure that's specific to that?

So for example, in the population of mental illness and addictions we want a specific smoking measure there if they can come up with one. So 40 some percent of the cigarettes smoked, right? Matt.

MEMBER STIEFEL: It would seem like a useful question to ask developers, if you're proposing that this measure could be

used in multiple venues or for multiple sub-populations, it's likely that the proposer is going to have to talk about the performance statistics and psychometrics for the population that they know.

And then make some assumptions about, we think it would also work for other populations. But we don't know, we don't have the performance statistics for application in a clinical setting. But it's at least worth considering.

MEMBER SAMPSEL: I think the only other thing I was going to add was, Helen, when you read back that question, it was not only about smoking, but tobacco use.

And so, you know, that's something we've been struggling with as we've been trying to do population health estimates of health of any state that WellPoint is in, that you may be seeing smoking rates go down, but you're seeing tobacco use and other mechanisms going up.

And so kind of the question to me is, are they also harmonizing on that line as well, that it's a dual question of both smoking and tobacco use.

CO-CHAIR JARRIS: I visited one --

MEMBER KINSINGER: I think that's a really important question not only for adults, but when you get to the under 18 population that's where you're seeing some of the major growth is in other forms of tobacco being used.

CO-CHAIR JARRIS: I visited one jurisdiction of each of the new health commissioners when they got put in place, where their tobacco smoking rate was extremely low. And it's because there's very high rates of marijuana smoking.

MEMBER KINSINGER: So you want to just ask if they're smoking. You don't want to know what they're smoking.

CO-CHAIR JARRIS: Okay, so --

MEMBER PESTRONK: Where the Youth

Risk Behavior Survey is a better one. Because it's asking about multiple --

MEMBER SAMPSEL: Cigars, cigarettes --

MEMBER PESTRONK: Cigars, cigarettes, legal, illegal drugs.

CO-CHAIR JARRIS: So should we open up now to public comment? Anyone in the room would like to make a comment?

MS. SHAH: Can I make a comment for Peggy, who's not here?

MS. MUNTHALI: Sure. Could you come up to a mic please?

MS. SHAH: She should be here.

CO-CHAIR JARRIS: Could you come to the mic so people on the phone can hear? And then we should try to rescue Peggy.

MS. SHAH: My name is Amishi. I work with Peggy. We spoke with Elisa yesterday on the phone. And one of our questions was that some of the measures that are already there, they do seem to be fitting

with public health quality aims. But they were not mentioned.

And one of our thoughts was, if people are really aware of what the aims are, and if there was a way that -- I'm going to lose my voice. If there is a way to make them aware of what the aims are, or at least let them know that these are there in there?

In the guidance document that is there if we could put a list of the nine aims or have the URL or something. This was one of the things that we had in mind.

MEMBER PESTRONK: Can I ask a question about that? Would you argue that, or suggest that all nine of the aims need to be met? Or that some threshold number of the aims need to be met in order to use the public health quality framework as an evaluative framework?

MS. SHAH: I would let Peggy have her final say on this. But I think that if we are looking at individual measures, I don't

think it may probably be possible to have all nine of them.

But if there is like, as you're saying, like a certain threshold, then that would be good. Because I think if we saw the adult HIV one and the one before that, which was -- Just the one before that? No.

The third one that we discussed. That was also something that we thought that it met the criteria for the aims. So even if certain are met I would be fine. But again, I think Peggy would have the final say on that.

CO-CHAIR JARRIS: That might be difficult in looking at them now, for a single measure to meet all of them. But --

MS. SHAH: Yes.

MEMBER BIALEK: I think demonstrating how they're considered, just discussing them could be sufficient to inform the committee's work.

CO-CHAIR JARRIS: Are people

familiar with them? You want me to just tell you what they are? Or are you familiar with them? That's not a yes or a no. I'm just getting blind stares back.

Okay. So it's population centered. I won't go into definitions. Equitable, proactive, health promoting, risk reducing, vigilant, transparency, effective and efficient. And you just google public health aims and you'll come up with them.

Thank you. Are there other? I guess Peggy's -- I asked and they didn't know where Peggy was outside the room. So I guess we'll give her a chance later to speak. Anyone on the phone like to make some comment?

OPERATOR: At this time to ask a question press star then the number one on your telephone keypad. At this time there are no further questions.

CO-CHAIR JARRIS: Okay. We have a 45 minute break for lunch scheduled. We are half an hour early, which I don't think anyone

will object to.

So why don't we shoot for 12, 1:15 p.m. to reconvene, if that works for folks? And people on the phone, just to let you know, that's when we all start to get back together.

DR. BURSTIN: The developers are welcome to eat lunch with us, please.

CO-CHAIR JARRIS: Thank you.

(Whereupon, the above-entitled matter went off the record at 12:27 p.m. and resumed at 1:16 p.m.)

CO-CHAIR JARRIS: Could we get ready to open the phone lines, please?

OPERATOR: All lines are open.

CO-CHAIR JARRIS: Okay. Why don't we open the lines, Operator.

OPERATOR: Did you want all lines open?

DR. BURSTIN: They're all open.

CO-CHAIR JARRIS: They're all open. Okay. Well very good. Well welcome back from lunch everyone. We had an

excellent, healthy lunch here, right? Following the NQF policy for healthy foods.

So the next step for us will be to review the committee recommendations for additional candidate measures clustered around BMI. And do we have representatives from the organizations available?

MS. MUNTHALI: Yes, we do. The first measure is measure 0029, counseling on physical activity in older adults. It's two parts. A is discussion physical activity and B is advising physical activity.

This is an endorsed measure by NCQA. Dawn, I think your colleagues are going to introduce the measure. And so, I'm sorry, I don't have your names.

CO-CHAIR JARRIS: If you could introduce yourself, please?

MS. GIOVANNETTI: No problem. I'm Erin Giovanetti. I'm a research scientist at NCQA. And I'm going to be presenting the measure 0029 on counseling on physical

activity in older adults.

So this is a survey measure that is administered through our health outcome survey. It's strictly for older adults, age 65 plus. There are two rates in this measure that are assessed through two questions.

The first question asks, in the past 12 months did you talk with a doctor or other health provider about your level of exercise or physical activity? And then gives some examples.

And then the second question asks, in the past 12 months did a doctor or other health provider advise you to start, increase or maintain your level of excessive or physical activity? And then gives some examples.

This measure is based off of the 4As model of ask, assess, advise and assist. Similar to the way that we designed our smoking survey measures.

The measure has good reliability,

which I think is all documented in the submission form, based on guidelines on U.S. Preventive Services Task Force recommendations about physical activity in older adults, and the importance of that.

I am happy to answer any questions. One thing that has come up before when we've presented survey measures to steering committees, just because we don't exactly know where to put it in the form.

Health outcome survey is a survey that is broadly fielded to Medicare beneficiaries. It does allow for a proxy response. So if someone is unable to answer this question for themselves, there is an option for a proxy response on this.

And there's a whole method for determining if someone is cognitively impaired enough to require a proxy response. Or the person can choose to have a proxy assisted response.

That may not be in the submission

form just because we weren't entirely clear where that would go in the form. But also happy to answer any questions.

CO-CHAIR JARRIS: Kurt.

CO-CHAIR STANGE: It would help me to have you make the case for why this is a population measure. I mean, the link between counseling by health care professional and actual behavior change is there. But I mean, it's probably maybe good enough to be a clinical measure.

But multiple determinants of health behavior change for this seems like the influence of counseling. Clinician counseling is just a really small factor compared to the others. So why is this not a clinical measure? Why is this a population measure?

MS. GIOVANNETTI: I guess I'm not entirely clear. This is a clinical performance measure.

DR. BURSTIN: Let me just speak to that. We included these remaining BMI and

weight measures in this part of the project. Because we were expecting to get population level measures that reflected that. These are clinical measures up for endorsement maintenance.

The question would be, what are -- As we go through this exercise, what are potential other population level measures that would be complimentary. But there is no expectation --

These measures are already submitted. We just moved them here because we thought we'd have population level obesity measures.

CO-CHAIR STANGE: Were they submitted in response to this call for population --

DR. BURSTIN: No. The prior work.

CO-CHAIR STANGE: So this is the old phase when we're still keeping --

DR. BURSTIN: This is the remainder of the measures from Phase I, where

we though we would actually have some very nice complimentary measures. We fully expected in this are that we thought we'd need to do some harmonization on. So we kept them in the project.

MEMBER PESTRONK: So you're asking us to endorse them, or adopt, or review them in the same way that we've reviewed the other ones?

CO-CHAIR JARRIS: Well we, in our criteria -- This is a population health measures work group. Do we have a criteria which would be such that we would only adopt a measure if it were a population health measure?

Because if it's not a population health measure, should we refer it to some other committee, whose jurisdiction it falls within? Or are we saying this is a Phase I, and therefore --

Because I could see that it is -- If you are looking at the 5,000 patients that

belong to health plan X and are surveying them, you're getting a sense of how health plan X's immediate population.

So really it goes back to Phase I. But this is not a Phase II issue. Prior to this committee, where would a question like this have been reviewed in NQF?

DR. BURSTIN: Probably just a prevention committee, or something along those lines. Or potentially behavioral health. We just really did expect we could get multiple --

You know, we specifically highlighted obesity in our call for measures for this phase. We thought there would be some logic of having the together. We're very surprised. And after lots of begging and pleading, got no population level measures around obesity.

Although interestingly, one of the developers for one of these measures is the City of New York Department of Mental Health

Hygiene.

CO-CHAIR JARRIS: Okay. So let's look at it within that light, that it is a patient population measure, which is Phase I of preventative service, clinical preventive service type measure. So, Sarah, you I believe from the committee reviewed this. Do you want to make any comments?

DR. BURSTIN: She's not here. But there were other --

MS. MUNTHALI: I'm not following my own rules. It was Rhonda Medows. She couldn't joint us today. But I think there were others in the workgroup that reviewed this measure as well. So if you'd like to share anything with your colleagues, please do so.

CO-CHAIR JARRIS: Okay. Anyone else have a comment or question? And again, let's hone in down and look at it from that point of view.

MEMBER KINSINGER: Paul, this is

Linda on the phone.

CO-CHAIR JARRIS: Go ahead.

MEMBER KINSINGER: And I wanted to make a comment again about the evidence that's quoted, which is all about the benefits of physical activity for older folks, which is completely non-controversial, totally agree, no problem.

But it would seem to me that the evidence that should have been presented for this measure was something -- Or at least commented on somewhere. Something about the effectiveness of counseling older patients about physical activity in the healthcare setting.

So again, and this has come up on several of them. I don't think the evidence that's presented is really the useful evidence that we need to judge the measure on. Just a comment.

CO-CHAIR JARRIS: Madeline.

MEMBER NAEGLE: I would agree with

that. And I just would suggest that since this measure was first approved there has been considerably more evidence about counseling about physical activity, in terms of levels and specific advice to particular issues. So I feel that the evidence which we have before us doesn't really speak to provider activities and their effectiveness.

CO-CHAIR JARRIS: And the evidence that you're aware of, would it support this measure or not? Are you prepared to talk about that?

MEMBER NAEGLE: The evidence that I'm aware of, and I guess I have to put on my other hat. That I direct our WHO Collaborating Center on Geriatric Education.

And a lot of the work that we've looked at really suggests that the advice about physical activity has to be targeted toward people's particular conditions, toward their capacity, toward their rankings in groups.

So that first of all, discussing doesn't really do anything. It's not paralleled to the 5As. I know that that might have been the original thinking about it now. But the evidence would suggest that that would not be appropriate as a parallel.

So that I think that it could be far more specific, and therefore more useful to both providers and consumers, in terms of what the consumer would expect to hear. And to the provider in terms of what the provider should be offering.

CO-CHAIR JARRIS: All right. At some point I guess then, if you want to follow up to that. And I guess the question is, would the instrument you're using allow you to get more specific in this fashion.

So that you could document that a useful intervention actually occurred. Because you're asking a patient. And would it work to ask a patient about that specific type of counseling?

MS. GIOVANNETTI: So that is, I agree with you. And I apologize that the evidence didn't speak to exactly about the counseling. We were really focused more around the guidelines around the activity.

We did look at more specific questions that really got at targeting whether or not they were targeted to the condition, to the exact circumstances of the patient.

We are constrained however, in that this is a large survey in which we had the option of two questions. Only two questions was the space with which we were able to do this. Because of the length of the overall survey, the health outcome survey, which gets multiple things.

So having a multi question tool would have been, could get at much more specific information. And I think is still something that is useful.

However, we were really looking for a measure that said across the entire

population of older adults, just was this even discussed in the physician's office?

And they did look at several questions that talked about more specific interventions. Patients found those confusing.

So in the cognitive testing, what they found was that it was better if you had the question be as broad as possible to include physical activity and exercise. Because people qualified those differently.

To not specify what the condition was, because for different people with different conditions, there might be different recommendations from their physicians. So this is really trying to get at the fact of --

It is a low level measure. I agree. It's not getting at what you want. It's not a guideline or any sort of advice to physicians about what they should be doing.

However, I will point out that even at this low level, we're still seeing a

pretty significant quality gap, with only 50 percent of patients reporting that their doctor even asked them about their level of physical activity. So I think it's still pointing to something that is a quality of care gap.

MEMBER STIEFEL: Two things, I guess. The first is, I'm drawing increasingly unclear about the distinction between a population measure and a clinical measure. And it sounds like you've got a bright line in your mind. But they're getting mushed together for me.

And perhaps that's a good thing. Maybe that's a direction we need to be going. This is a, you know, you could argue that this is a sub-population measure, if you're looking at populations of health plans.

And I'm not sure of this in or out notion of only geographic population applies. I agree this is a low level. This is a major strategic initiative in our organization, to

be doing this measurement. And including measuring exercise is a vital sign.

Perhaps the efficacy of the advice at the clinic visit is not as great as other interventions. But promoting awareness of this is I think powerfully important. And this is a step in the right direction, like a lot of the other measures we're looking at where the measure itself --

(Off the record telephone interruption.) CO-CHAIR JARRIS: Okay.

MEMBER STIEFEL: What was I saying? That this is an important component of a broader -- Like, you know, smoking itself isn't a population improvement measure. Assessment of physical activity is a first step in a very important population health set of measures.

CO-CHAIR STANGE: This isn't an assessment of physical activity, right? It's clinician counseling for people 65 and older. And to Madeline's point about the evidence for

the efficacy of counseling, which I think is relevant to this.

I mean, I -- There really has been a lot of research done that's come out since the task force looked at this. Have they looked at this again since 2002, which is what I just googled?

MEMBER SPANGLER: Yes.

CO-CHAIR STANGE: So what's the latest recommendation?

MEMBER KINSINGER: It will be coming out in about a month or so.

MEMBER SPANGLER: Yes.

CO-CHAIR STANGE: Okay. So --

MEMBER SPANGLER: And did you have a systematic review that came out in December of 2010, hereabout? Behavior counseling to promote physical activity, and a healthful diet to prevent cardiovascular disease in adults. But not beside that, and in that specific population, elderly population.

CO-CHAIR STANGE: So even, I mean

--

MEMBER SPANGLER: It's adult.

CO-CHAIR STANGE: Even if we're able to extrapolate to the older populations, we so often have to do, does the evidence produced support Madeline's contention, and my impression that there is evidence, a greater body of evidence of the efficacy of counseling?

MEMBER SPANGLER: The conclusion from the systematic review is counseling to improve diet or increase physical activity, change health behaviors, and was associated with small improvements in adiposity, blood pressure and lipid levels.

CO-CHAIR STANGE: So for this measure that seems very important. It's not really about just changing the behavior, rather does counseling make a difference. Okay.

MEMBER STIEFEL: Paul, could you respond to the question that I raised about

where is the line --

CO-CHAIR JARRIS: I would put, you know, if we look at the NPP. And I think you may have been the one that actually came up with the three tiers originally. I would put this --

This is a clinical behavior, a clinical behavior service. One on one type issue. In that case I would put it in that Tier I, looking at clinical preventative services at a population level.

In this case the population is the health plan, or whoever is using this. I think it doesn't get -- And I would agree with you, that's necessary, but wholly insufficient intervention in the whole system of interventions.

To go beyond that, it would be great to have a measure among, of the actual physical activity level. For example, in a jurisdiction or a population. Because then it's not a matter of I counseled you, it's

actually a step beyond that.

Not only did I talk to you about it. But I have to make sure that there's a safe place to walk in your community, that you can go out after dark if that's when you exercise, without being assaulted, that you can go to playgrounds. So it gets beyond the four walls. This is still a within the four walls question.

MEMBER STIEFEL: We just talked about smoking though. And it was just an assessment of do you smoke?

CO-CHAIR JARRIS: Which is probably the same category here.

MEMBER STOTO: No. I think it's not. If it was --

CO-CHAIR STANGE: That was a different measure, you're right. That was prevalence of smoking in the population.

CO-CHAIR JARRIS: It's about do you smoke, which is different than did someone counsel you about smoking.

MEMBER STOTO: I think these are really fundamental issues that the fact that we're struggling with these strikes mean that the people developing the measures probably are struggling even more.

And maybe we should put this off until tomorrow. But I want to raise a flag here that I think would be useful to provide some guidance with examples, perhaps. That sort of say, look at these different kind of measures.

How do they fit into the population health logic model? What kind of evidence do you have to propose for each one? And things like that. I think that could be useful.

MEMBER PESTRONK: This is a proposed measure that's coming back after having been approved, for maintenance. So is there anything that has happened in the last three years, roughly, that would suggest that it shouldn't be endorsed again?

Has anybody raised issues to suggest that, you know, this one didn't work? Or we've got problems with it? Because somebody did a thorough review of this the last time in another place. So what do we know about that?

CO-CHAIR JARRIS: Well from what Jason said, the review was in December. And it supported that there's some level of evidence that this probably does some good.

I mean, we could also -- If this kind of assessment and counseling isn't going on, that's bad. If it is going on, it may be good.

So from that point of view, if 50 percent of the folks are responding that they never even were assessed and counseled, that's bad. So I do think there's a gap here that could be improved.

It's just with this, again I think we have to be clear that this is like the Phase I of the three phases. And not get lost

and think we've done our job by approving these very narrow, in the four walls type, measures.

MEMBER PESTRONK: That I get. I've got a brighter line in my head between, you know, measures like this and measures that we have tried to conceive as population based measures.

And the smoking is an example of one of those that seemed to fall way over on the other side of the bright line, on the brighter side of the line.

The question for me, setting aside all that, in trying to answer the question that we're getting asked, is should this be re-recommended? Has anything happened in the last couple of years that suggests that it shouldn't be?

And if nothing has come up that it shouldn't be, and people are using it, and they're getting good value out of it, and there's variation in what's being done, and

therefore being able to have this measure illuminates that variation, and moves people towards reducing the variation, then --

DR. BURSTIN: And just to reflect on that. I mean, when a measure comes back in for maintenance, it's not just that nothing bad has happened as a result of having the measure out there. But has anything positive happened as a result of having the measure out there.

So perhaps NCQA could speak to -- I mean, it looks like at least the section 2B5 indicates that across Medicare plans there's been little changed. Now again, it's not clear how much of that, you know, rates across the years have not changed much.

But maybe that, you know, in terms of future tense, that may point to Paul's earlier point. That if it's only a discussion between you and your provider -- And I'm curious about Sarah's perspective on this as well. As opposed to your health plan.

Potentially doing activities outside of those four walls, which may get you closer to more of a population level measure. Just an interesting discussion I think for us to have. But I think we're really seeing not much difference. Erin, do you want to respond to changes?

MS. GIOVANNETTI: Yes. So there was little attention paid to this measure. I agree. Which is why you see in the data that we have up until 2009 reported here. You don't see much change.

However, it has been adopted into the Medicare Stars program, Stars rating program, which has put a lot more attention on these measures. And so we are --

And unfortunately, you know, just because of the data delay we don't have those numbers directly available right now. But we do expect that with that, this measure will start to move. Because if anything, people want to increase their stars.

CO-CHAIR JARRIS: Elisa, did you want to address this issue of --

MS. MUNTHALI: Yes. There was a question about whether or not we received any information on its use. And what we do as part of the endorsement maintenance process --

When we put out the call we ask people in the field if they have any comments about the measure. And just as a matter of public record, we just wanted to note one comment that we received. And we shared it with NCQA. I don't have your response. So if you want to respond verbally you can do that.

This comment is from Carmella Bocchino. And she's one of our member counsels. So she recommends that NQF harmonize this measure with language used in the HOS survey, particularly -- This is in quotes "physician recommending an increase or maintaining exercise."

Additionally we recommend NQF explore additional means of reporting, beyond

self reporting. Since exercise can be categorized as a vital sign.

And so we just wanted to pose it to you to respond to that. I think, Dawn, you sent us your response. And we're trying to find it on the P drive.

DR. REHM: Yes, hi. I'm Bob Rehm. I'm the Assistant Vice President for Performance Measurement at NCQA, and formerly worked with Carmella at AHIP.

And actually had a conversation with Aparna Higgins, who's a member liaison to our committee on performance measurement, about the perfectly harmonized measure within a measure.

And she said, yes, that was an error on our part. So it is the measure that it is, which is the one that's self reflecting I guess. I don't know.

(Off microphone comments.)

MEMBER STOTO: So I've got just a technical issue. the denominator statement

actually is a statement of the numerator, if you read it. It's 2A1.4.

And you can infer what the denominator is. It's just contained in there, Medicare members of 65 and older. But it then restates to numerator. So just to be clear.

MS. GIOVANNETTI: That was something that got messed up in the transmission of the form. I apologize. I don't know what happened. But that's not the --

MEMBER STOTO: Oh, yes. Well not quite, but -- Anyway, that's a technical thing that --

MS. GIOVANNETTI: Oh, no. I apologize, no. The reason that that is there is because the answer to the question is yes, no, or I've had not physician visit in the past 12 months.

Anybody who reports yes or no is in the denominator. If they report not having had a visit in the last 12 months, they're not

in the denominator.

MEMBER STOTO: Oh.

MS. GIOVANNETTI: So we only include people who had a visit in the last 12 months.

MEMBER STOTO: I see. I didn't get that. But okay.

CO-CHAIR JARRIS: This is a four walls question then. Because you're not asking the entire plan's enrollment. You're just asking those who had a visit.

MEMBER STOTO: And who answered the question.

CO-CHAIR JARRIS: Madeline, did you --

MEMBER NAEGLE: I would just ask that we might clarify if we're going to say this has worked so far. But you're still not getting great response. I think it was about 50 percent are actually doing this.

Rather than using vague terms like discussing physical activity -- You used the

language, Paul, when you were just talking about it. Saying did you assess and counsel? So maybe it should be assessing physical activity and then advising. And that would maybe be a little bit more specific.

CO-CHAIR JARRIS: Yes. Someone has to translate that into patient language.

MS. GIOVANNETTI: And we did. We tested that language with the patient, with did the physician assess your -- Although we were not asking specifically about whether --

In the cognitive testing I don't think we were getting at did they assess your level of physical activity. That, we were more asking about did the doctor ask you about it.

And they did test a variety of different languages, which is why they ended up on -- Discussing was the term that made the most sense in this older population.

MEMBER STIEFEL: This just right to that point. It sounds as if what Paul and

Bobby were saying earlier is that this would be very different if this were a measure of assessing physical activity. Then it would look a lot like the smoking measure.

And this is just a little bit removed from that. It's discussing physical activity, which isn't quite at assessment. But it was, in part, because it was put in consumer friendly language. So it's close to an assessment measure.

We've added exercise as a vital sign. So we're measuring it as well at every visit. But it sounds like that's a very important distinction. Kind of a subtle distinction.

MEMBER PESTRONK: It's a sub -- To that point it's a sub-population, which is non geographic in nature. So if, in fact, we want to think about sub-populations that are non geographic in nature, then it has the potential to be a measure that has to do with the population, which is the clinical

population in a particular organization, or in a particular plan. That's the distinction that I was making in my own head on this one.

CO-CHAIR JARRIS: Does anyone else want to talk before we allow Ron to say something?

MEMBER BIALEK: I feel like I'm home. Well actually it's okay. Because I'm very confused at this point. You know, I think of some of the earlier comments were about impact.

And so we have counseling here. And we have physical activity here. And at least, again what I'm reading, that link between counseling leading to the physical activity is not made. And I know Jason added some information that was important. But I don't see that being made.

And so my question is, if we are to apply the same criteria that we did this morning, where we look at, let's say 1B, performance gap, demonstration of opportunity

for improvement in health.

It doesn't suggest to me that the counseling measure has demonstrated an opportunity to improve health, because the counseling measure doesn't make that tie to increased physical activity.

CO-CHAIR STANGE: So I think we've done two re-frames. At least since I was initially voting on this and thinking about it. One, is that we're thinking back to the clinical preventive service delivery measures, and not holding this to a population health standard.

And then second, I think is Bobby's very helpful comment that this has been previously endorsed. So what's new? The new thing is that now it's actually got some teeth, some incentivization for people to want to use it. That's important. That increased the impact.

And then Madeline and Jason talked about if anything, the evidence for what Ron

just brought up is certainly, probably very greatly enhanced since the last time it was looked at.

So to get back to Bobby's frame. What's new? The two things that are new are in a positive direction. It was already endorsed before.

So I'm not hearing any discussion about why we shouldn't re-endorse this with that initial kind of re-frame, that this is a clinical preventive service delivery measure.

CO-CHAIR JARRIS: Okay.

MEMBER PESTRONK: Does every measure that NQF has endorsed get used by someone? Or do the measures that NQF endorsed sit out there as a collection of measures to be chosen by someone to be used? And then, so that's one question.

The second is, progress in the direction that we would hope isn't the right question, isn't the right framing. Because failure on a measure is an indicator of what

shouldn't be done, if people have been trying to have progress occur in the right direction.

It suggests what people shouldn't try again. Or shouldn't try again in exactly the same context. So, Helen, that get's back to, you know, the question is, has this been chosen? And now it looks like it might be chosen.

And I don't care whether there's progress or not. What I care about is somebody pays attention to the analysis about why there was progress or not.

And that things don't get repeated over time. People don't do the same things over and over again, failing to learn from the experiment that they conducted.

CO-CHAIR STANGE: So just to be clear, Bobby, I wasn't talking about progress in actually the health behavior that's being measured, or the counseling. Progress in meeting the criteria is what I was speaking to --

MEMBER PESTRONK: Okay.

CO-CHAIR STANGE: -- with your initial re-frame.

DR. BURSTIN: But does it answer your question? We actually have been trying to track uptake of NQF endorsed measures. And in fact this past year did a very detailed analysis of all 700 measures in the portfolio.

And the overwhelming majority -- I think it was almost 90 percent, are in use in some program. In fact, a good number of them are used in federal programs, not surprisingly.

A good number of them are used by states. Some of the state based measures from AHRQ, for example, the RQIs and others, PQIs. And then there's a set of measures that are used in benchmarking, quality improvement, a series of things.

Not all of them have moved into sort of highs stakes accountability applications. But the vast majority of them

are in use.

We've also learned, as part of our usability task force this past year, it usually takes at least one cycle of a measure being out, a new measure, to get really picked up for use.

MEMBER STIEFEL: Perhaps this is just a tomorrow question. But the point that Bobby raised is an important one where I might disagree.

The notion that if we're looking at a non-geographic sub-population, does that mean we can't have population health measures? Because a lot of what I do is population health measurement for a non-geographic sub-population.

And would argue that if we could find harmonization between Kaiser and WellPoint and United, and the public health department in a geographic area, the union of those sub-populations comes close to being the population. And I think that's a very

important direction.

MEMBER STOTO: Can I just add to that? I mean, I agree with what Matt said. But I also think it's important to have these population health measures for the population you said to begin with.

Because so for instance, take something like the immunization measures we were looking at last time. If you measure immunization status among the people who happen to have come in during a given time period, versus immunization among everybody who's enrolled in the program, they are fundamentally different. And they tell you different things.

And I think that the second one is the population health measure. And I think that's the one that we should be doing, even for service oriented populations.

CO-CHAIR JARRIS: Okay. Any of our committee people on the phone have something to say? Otherwise I think we're

probably ready to move on.

OPERATOR: To ask a question press star then the number one on your telephone keypad.

CO-CHAIR JARRIS: Actually, I wasn't opening up to the general population. I just thought if there were some committee work members. Maryann and Linda?

MEMBER KINSINGER: No further questions from me, thanks.

MS. JACOBSON: This is Dawn. Can I make one comment as a public person?

CO-CHAIR JARRIS: Go ahead.

MS. JACOBSON: I just want to reiterate, based on what we found in the environmental scans, that the measure is currently submitted as a health improvement activity of the clinical care system. It's not a population health measure.

The solution would be using something like BRFSS to measure physical activity rates at the population level. In

the future, ideally a health information exchange that could pool all of that information, and be used geographically to assess counseling at the population level would be another solution. But we don't have that in most areas of the country yet.

CO-CHAIR JARRIS: That raises another possibility. Maybe this is another one of those areas where it would be useful to have, what is the question. And have it used by CMS and have it used by Kaiser, and have it used by BRFSS, so it was harmonized. Now what that has to do with what we're doing today confuses me.

MS. JACOBSON: I would just recommend that we use a measure like prevalence of physical activity, using an existing survey. Because our limitation right now is we rely on these self reported population health surveys to make decisions for change.

So I don't see this measure being

a population health measure. And, you know, Healthy People has all sorts of population health measures that have been vetted for similar things, like validity, accuracy, what data set do we use, what's meaningful at the local level.

And I'd like to see measures that are truly population health measures be selected, not clinical care measures. I guess is -- I don't if that means we have to redo it.

CO-CHAIR JARRIS: Yes. This would be more of a population health measure if the question was, are you, whatever it happens to be, are you physically active 30 minutes, three times a week, or whatever the question is.

And then it would reflect upon that population served by that insurer. That's more of a population measure, as opposed to were you counseled, which is a clinical intervention measure.

MS. JACOBSON: Yes.

MEMBER STOTO: Can I disagree with that? I mean, I think that -- Was that Dawn speaking? I think that that reflects the distinction they make, and she makes in the paper about population health outcome measures, as opposed to population measures.

Because I think that you can talk about doing an activity like counseling among everybody in a population, geographically or defined as Matt says. And that's not -- And that's of use.

That's distinct from an outcome measure. And this one's about the counseling and the discussion, rather than the physical activity as an outcome.

MS. JACOBSON: Right. And maybe we can talk about this tomorrow again as well. But if we had a valid, reliable population health data set that gave us that information, we could then logically use it as a proxy for downstream outcomes if the evidence clearly

links it, which I don't know if it does.

But we're limited with our data sources right now, I think. And I wouldn't use a health plan data set to assess population health the way it's collected no.

CO-CHAIR JARRIS: All right. So there's a lot to discuss tomorrow. I throw into that what's also the question, if you have actually population health outcome measure, would you necessarily use that rather than a process measure like counseling. For tomorrow maybe.

DR. BURSTIN: In some ways those may be complimentary. I have a question for NCQA, since you guys are here, in terms of question wording. And this came up earlier before you came in. Dawn was helping us out with looking at the smoking measures.

Do you look towards the national surveys for the wording? For example, BRFSS or NIHS to determine what questions you'll ask?

Because one of the questions that's been asked is, if you want to get a family of measures, you'd certainly like to have the same word choices used for your surveys at the health plan level as we're seeing at the population level.

MS. GIOVANNETTI: So I can answer to the, you know -- Definitely as we develop new measures we look to already established measures questions. We definitely want to go with --

If there's something out there that's validated, that gets at what we want, we go with that. When this measure was developed, back in 2003, we did the cognitive testing in 2004, I wasn't part of the team that developed this.

But that's part of the environmental scan that they do. They look across measures that are existing to determine if there's one that they could use and put into that survey.

I suspect for this one, there was not one existing at the time that they felt was appropriate for the 65 plus to be administered. It's a self administered survey.

DR. REHM: And just to the smoking cessation question. Originally when NCQA promulgated that measure, it was using the very early public health service ask and advise to quit. There was the two questions.

And then as the public health service guidelines expanded, I think Dr. Fiori led that, then we added to it essentially. So this is a measure that essentially grew over time versus a essentially as a place holder for a significant public health threat.

And we had, you know, a lot of resistance among different communities who didn't want us to go there. And so, I mean, some of these measures have a life story. And smoking is one case.

This is, of course, an area where

as we get new recommendations from the USPSTF, coming soon, we would imagine we would look forward to the opportunity to make sure that we are capturing all those elements that we could.

At least in this area within the constraints of the health outcome survey itself, which is a fairly burdensome survey as it stands. But again it is an instrument of accountability, which we think is quite important.

CO-CHAIR JARRIS: Okay. We should move on and I think probably to the voting process. Yes, why don't we first -- Sue, I apologize that we haven't given you an opportunity to introduce yourself to everybody, let the folks on the phone know you're here. And we also updated any conflicts. So if you could do that.

MEMBER PICKENS: Sue Pickens, Director of population medicine, Parkland Health and Hospital System, Dallas, Texas. I

don't believe I have any conflicts at this time.

CO-CHAIR JARRIS: Thank you. Elisa.

MS. MUNTHALI: Okay. So we'll start with importance to measure and report. 1A, impact. The rating scale, just to remind you is one for high, two for moderate, three for low and four for insufficient evidence.

And remember, Sue you wouldn't remember this, but make sure you point, yes, your clicker to this monitor here. And we have 20 seconds to record the vote. So the timer is on. We need three more votes. And one more.

(Off microphone comments.)

MS. MUNTHALI: Okay. So Mary and Linda on the phone.

MEMBER KINSINGER: This is Linda. And I would say moderate.

MS. MUNTHALI: Mary Pittman, are you still on the line?

MS. MUNTHALI: Okay. So there's seven highs, three moderates and one low. So the next category is 1B, performance gap. Same rating scale. You have 20 seconds. I think we have five more in the room. I can't see what number that is. Two more. One more. Okay. We're fine. Linda?

MEMBER KINSINGER: High.

MS. MUNTHALI: Okay. So ten highs and one insufficient. So we'll move on to 1C, which is evidence. And the rating scale is one for yes, two for no, and three for insufficient evidence. Okay. We have everyone in the room. Linda?

MEMBER KINSINGER: Well based on what was submitted, I guess I'd have to say insufficient.

MS. MUNTHALI: Okay. So we have six yes's, three no's and two insufficient. So we move on to scientific acceptability of the measure properties.

This is the evaluation of

reliability. Rating scale from high to insufficient evidence. This is 2A. I think we're missing one. Okay, we have everyone in the room. Linda?

MEMBER KINSINGER: High.

MS. MUNTHALI: Okay. So seven highs, three moderate, one low. Next is validity, 2B. Same rating scale. We have everyone in the room. Linda?

MEMBER KINSINGER: Moderate.

MS. MUNTHALI: Okay. So six high, four moderate, and one low. So we'll move on to usability. Same rating scale. We're missing one person in the room. Okay, we have everyone. Linda?

MEMBER KINSINGER: High.

MS. MUNTHALI: Okay. Seven high, three moderate and one low. And the last criterion is feasibility. Same rating scale. We're missing one. Okay. Linda?

MEMBER KINSINGER: Moderate.

MS. MUNTHALI: Okay. So three

high and eight moderate. And so now we're going to vote on the overall suitability for endorsement. Ths is a yes or no vote. One is yes, two is no. Okay. Linda?

MEMBER KINSINGER: Yes.

MS. MUNTHALI: Okay. So ten yes's and one no. So the measure passes.

CO-CHAIR JARRIS: Thank you. As you can tell we're having some adjustment issues here with the new work area. So that reflects us and not you. And thanks for your work.

Please consider what we talked about though in terms of how to make a population more broader population measure, as opposed to an encounter measure as you go forward with these things. Thanks.

Next we are back up to New York City. Do we have folks on the phone from New York who'd like to speak to Measure 0023, BMI in adults older than 18 years of age? And I just ate at Carnegie's Deli this weekend. So

that's where you have to start. One sandwich took four days to finish.

MS. MUNTHALI: Operator? Anika?

OPERATOR: Yes.

MS. MUNTHALI: Hi. Is Sam Amifar on the phone?

OPERATOR: Hold a moment.

MS. MUNTHALI: Okay.

OPERATOR: If you're on line can you please press star one? And their line is open

MR. AMIFAR: Hello?

MS. MUNTHALI: Sam?

MR. AMIFAR: Hi.

MS. MUNTHALI: Hi. This is Elisa. If you could go ahead and introduce the measure for the committee, that would be great.

MR. AMIFAR: Okay.

MS. MUNTHALI: Thank you.

MR. AMIFAR: Sure. I'm Sam Amifar from the New York City Department of Health

and Mental Hygiene. And the measure that we put in front of you, I think the number was 0023, Body Mass Index, BMI in adults greater than 18 years of age.

CO-CHAIR JARRIS: Yes. That's correct. Would you like to give us any background on it?

MR. AMIFAR: Sure. So this measure was developed in about 2006. And the purpose of it was to measure a number of BMI measurements that providers were taking.

The suspicion was that actual numbers were pretty low, in terms of height and weight being taken in a lot of our clinics. We wanted a measure that would give an administrative associate or manager of the clinic an idea of the rate of BMI being recorded.

At the same time have a measure that on the alerting side would alert the provider that a particular patient in front of them being seen, would need to have height and

weight measured so that BMI can be reported for that particular patient.

The idea was we would have probably a multi-stage process with the measure. Ideally down the road we would like a measure where if a BMI is high or out of range, either, you know, obese or extremely low, that a measure would prompt the provider to do something about it. Counseling and discussing options for an obese patient for example.

The first step of it in 2006 was however, to get a baseline recording of how much BMI is being recorded. Then the measure was incorporated into a lot of the practices. Bu 2007 we got our numbers back. And I think we had about 20 percent compliance rate with the BMI.

And over the course of the 2008-2009 the numbers steadily improved. I think the last numbers we have about 65 or 70 percent. They might even be higher now. We

don't have current years, 2011 numbers yet.

And the idea was once it gets into a certain threshold, we're thinking maybe 80, 85 compliance with it, then we would modify the measure and make it not only measuring BMI, but also working on the intervention part of it.

CO-CHAIR JARRIS: How was it -- Is this a -- How was it submitted to you? And what sense do you have of how much of the issue is that it's not done versus how much of issue is that it's done but not reported to you?

MR. AMIFAR: It's submitted -- We have about 600, 700 practices, about 2500 providers now. The practices have their own servers. And the electronic health record submits it to us electronically and automatically on a monthly basis. And is sent to our servers at Department of Health for analysis.

And so we get these automatically.

And the provider doesn't have to do much in terms of -- Actually it doesn't have to do anything in terms of having these numbers sent to us.

There are occasional transmission issues. And some practices do not transmit, and haven't been transmitting because of technical issues with the transmission.

But for the most part we get a good representation of our different practices. And could you repeat the second part of the question again?

CO-CHAIR JARRIS: No. You answered it. Thanks.

MR. AMIFAR: I did. Okay.

CO-CHAIR JARRIS: Thank you. Who among the committee reviewed the measure? Jason?

MEMBER SPANGLER: Yes. That was me. My response is, I don't know if you saw the memo. Most of them were in there. Just some questions about, some of the responses

didn't seem adequate enough. And I felt like some of the citations were just a little weak.

There also, one thing to note is that, this is also a topic in progress for the task force. And they are in the final stages now too of updating their screening and management of obesity in adults. So that's something to keep in mind. It should be coming out in the next few months.

One of the other questions that I had was the time frame, the 24 months. And the developer I guess responded to that, that there wasn't evidence for that. But I guess it was convenient for the providers there.

And then gave more details about the exclusions. So my main issue with this measure it that it's -- And I really appreciate this side by side comparison. Is that there are competing measures for this.

So two that are endorsed, including this one and 421. And then a new one that NCQA submitted, 1690. So I think we

need to address the competition and harmonization of those.

And they have slight differences, each one of these. And I think each of the elements have, I would say good components that we could probably --

I don't know if we would combine them or add to one of the measures that have the most. But I think we should probably, I mean, we should only have one measure I think that covers BMI.

CO-CHAIR STANGE: So based on Jason's comment that the U.S. Preventive Services Task Force is going to come out with new recommendations soon. And based on the measure developer's plan that this becomes a measure that includes counseling beyond measurement.

And based on the harmonization issue that Jason brought up. I mean, is this something where we ask the NQF to look at harmonization issues? And then updating based

on --

DR. BURSTIN: No. We ask the developers to harmonize.

CO-CHAIR STANGE: Okay. So we ask -- I guess the question is, if the task force comes out with something that's, and maybe people have some inside information on this, that's beyond just the screening, that's screening and counseling. Is that, if that's where you want to go, would you want to update this to reflect that?

DR. BURSTIN: It's the New York City's guys on the phone.

CO-CHAIR STANGE: Right. Right. Thank you. I'm sorry, so that was a question for the developer, the New York City Health Department. I'm sorry, I don't have your name.

MEMBER SPANGLER: Sorry. Kurt, can I address that real quick?

CO-CHAIR STANGE: Yes, thanks.

MEMBER SPANGLER: The old

guideline they use, the SPS evaluation does also include counseling. So one of the questions I had, which I don't -- Actually I don't think it's reflected in the memo.

But one of the questions I have written down on my -- And I think we brought up during the call was, why is this not a measure for both components? Just for screening and not for counseling?

DR. BURSTIN: Although I was actually at the task force at the time. And it was specifically, if you were in settings that had specialized obesity programs. So it was nuanced at the time. You could pull it up. I'm curious to see if there's better evidence now.

MEMBER SPANGLER: So that's a great question. I had that question as well.

CO-CHAIR STANGE: For the measure developer. Do you want to think about updating this based on what the new task force recommendation?

Particularly since that's your intent, is to move in the direction of a measure that's about screening and counseling? Recognizing that there still are many practices for whom the screening is a current unmet step.

CO-CHAIR JARRIS: Sam, are you maybe muted? Or we could repeat the question.

MR. AMIFAR: Hello? Can you hear me now?

CO-CHAIR JARRIS: Yes.

MR. AMIFAR: Okay, great. Sorry about that. So back in 2006 when we first introduced the BMI measure we didn't incorporate counseling. Primarily because the electronic health record vendors we were dealing with didn't have a mechanism of incorporating counseling sessions into electronic health records.

And so since we always knew that this was going to be a preliminary BMI, that we would want to do an intervention down the

road at sometime.

This was meant to be a stop gap type of measure, where we would try to increase the BMI recording rates to a certain respectable amount.

And then once that happened then we would want to update the measure. And in terms about updating the measure. We obviously are very interested in updating the measure.

We haven't updated this particular measure because, like you said, there are a couple of measures that are in this class and are competing.

And I think specifically 0421, if I have the number right, is a measure that we do like in terms of an annual BMI screening that incorporates counseling. And with BMI parameters of greater than 30, or less than 20 or 18.5 were numbers we were comfortable with.

We were thinking of incorporating that. And we didn't want to update this

measure and have one that's very similar to the one that's already being considered. And it's already endorsed.

CO-CHAIR JARRIS: So do you -- In looking at this then, would you be supportive of measure 0421?

MR. AMIFAR: We would, you know, in terms of I think having too many measures in the same field that are competing or are not very different from each other, I think is confusing for the public in general, the developers and practitioners.

So we would be supportive of 0421, or another one the group determines to be a good representative measure that's supported by evidence currently.

CO-CHAIR JARRIS: Mike and Sarah?

MEMBER STOTO: I think it's important to be careful about exactly what's being measured by these three things on the chart.

And that this, the one that we're

currently discussing, 0023, is similar in concept to 1690, which is about documenting BMI. But 0421 is about actually doing something when you find somebody is out of the normal range.

I mean, that's fundamentally very different. And I think that what Sam was saying earlier is that they're thinking about this 0023 as something they can do now, so that at some point in the future they could do something like 0421.

But it's not like one is a substitute for the other. No. I don't think it's a substitute at all. I think that they address very different things.

I mean, right now, I mean, I can imagine that if the physician measures the patient's height and weight. And then the system calculates the BMI. It may not ever tell the physician even what the BMI is, and whether the patient's out of range. And that's --

MEMBER SPANGLER: I think what you're -- Mike, what you're pointing out is, if you look at 0421, that and is capitalized. So you need both components. That's what you're talking about.

MEMBER STOTO: Right.

MEMBER SPANGLER: Not just the first one.

MEMBER STOTO: Right.

MEMBER SPANGLER: It's not an and/or.

CO-CHAIR JARRIS: Right. But isn't that what Sam is saying? They're ready to move on.

DR. BURSTIN: They're ready to move on.

CO-CHAIR JARRIS: As long as their electronic health system's ready to move on. So wouldn't we, you know, drive -- This reminds me of the diabetes one. The first question was, do you get an A1C once a year? And the question after a while is, duh?

MEMBER STOTO: Yes.

CO-CHAIR JARRIS: And then, what's the level?

MEMBER STOTO: I guess maybe we should ask Sam. Because I heard him say maybe that they were not ready to do something like 0421 yet.

DR. BURSTIN: I thought he said he was.

CO-CHAIR JARRIS: I thought he said he was too.

MR. AMIFAR: Hello?

CO-CHAIR JARRIS: Yes, Sam.

MEMBER STOTO: Okay, you're the one.

MR. AMIFAR: Okay. Great. We actually are ready to do something like 0421. We brought the preliminary measure up to about 75, 80 percent.

So I think once it gets that high we're ready for actually some kind of intervention. And we'd be supportive of

counseling involved in patients who have a higher BMI.

MEMBER STOTO: Okay. Thanks. I misunderstood that.

MR. AMIFAR: Okay. No problem.

CO-CHAIR JARRIS: Sarah, then Kurt.

MEMBER SAMPSEL: And I think this was touched upon a little bit in that part of the conversation. In that one of the things the workgroup was talking about is how do you deal with these multi-component measures.

And we'll deal with them again in a few minutes. But, you know, so if we're looking at 0421, while there's an and, there's also an if in there. And so, you know, it's a translation issue.

And so if New York's moving to that as well, that's something also to kind of consider. Because BMI, in and of itself is, you know, as Matt probably knows as well, is one of those things we're trying to get

doctors to document.

And so even alone it's a useful piece of information. In moving patients, you know, whether up in BMI, which doesn't seem to be a huge problem in our country, or down in BMI. You know, that's kind of another clinical process that we would have to get to.

But the other thing that I would wonder from New York would be, if you get to that point where you're going to that, you know, the 04210 where the doctor obviously has the BMI in front of them.

Will you be capturing that data too, so that you then be able to have population based metrics for New York on really where the population is falling on the BMI scale.

MR. AMIFAR: In terms of actually getting numbers out of the EHRs, we don't get actual values out of it. We do get categorical groupings. So we probably would get a count, for example. A monthly count of

people who have a BMI over 30.

But actual BMIs being sent to us on a per patient basis, we would not be getting. We generally just get counts of a certain categorical variable that we're interested in.

CO-CHAIR JARRIS: Sam, the clinics you talked about that are reporting this. Are those the New York Department of Health clinics? As opposed to all the providers and all the medical practices. Because you'd only be rolling up a subset of the population.

MR. AMIFAR: Right. These are actually not Department of Health clinics per se. They are practices that are associated with our regional intervention center. A primary care information project.

We, for the past five years have been helping practices, private practices, community health centers to adopt EHRs, and to start using them effectively. And so they're primarily small practices out in the

community, with a couple of community health centers.

CO-CHAIR JARRIS: But they're not necessarily representative of the whole New York population?

MR. AMIFAR: They are pretty representative of the New York City population. We have about 2.5 million patients represented in New York City. It's slightly predisposed to the underserved community.

Because under our initial grant that's who we were, those are the providers we were targeting, people who've had at least maybe 30 percent Medicaid patients.

So we are biased towards the underserved community. But we think overall it's a pretty good representative sample of the New York City population.

CO-CHAIR JARRIS: Kurt?

CO-CHAIR STANGE: So for Sam's population practices and his REC, the CMS

endorsed measure is the next step. The question I think for us is whether it's useful to have a measure that's more of a developmental measure for the practices that aren't there yet.

And I don't know what the national data are. And it's a rapidly evolving area. But I doubt that 79 percent of the practices nationally are reporting BMI. There's still a big lag for, first of all measuring height.

Second of all, for people with paper records BMIs tend to not be calculated. And there are still a number of electronic health records that surprisingly don't calculate BMI.

So the question is do we want, as a developmental stage for the next few years, while everybody's catching up with doing this, do we want a measure that just does the assessment, and doesn't include doing something with it?

I mean, is that a useful thing?

And does that fly in the face of NQF's desire to harmonize? Or is that really saying, the field is at a variable stage?

There's some for which this is too low a bar. And there's some where if you set the bar higher you don't give them the stepping stone to get there.

DR. BURSTIN: These are all good questions. But usually our process is, we go through each measure. We indicate whether it's suitable for endorsement. And then we have this exercise with the developers.

Because we have not actually heard yet from Quality Insides or NCQA about the measures they put forward, where they may actually have some suggestions about why their's is a better mouse trap.

But I guess my major question for New York is, do they feel they want to continue to be the -- To maintain this measure as an endorsed NQF measure?

And if not, would they be willing

to work with the other developers to kind of potentially works for their needs, the Board of Public Health needs, and the needs of the other developers here.

I mean, I will point out 0421, as I understand it, has been retooled for EHRs already, as part of meaningful use. So it might be not too much of a heavy lift for New York City to adopt it into their measure, which I believe is based on E-clinical works, one the EHR they adopted.

MR. AMIFAR: Yes. In terms of the New York City viewpoint. I mean we'd be happy to work on a new measure. I think, you know, New York City might be a little bit ahead of the curve in terms of nationally where we are. And we're ready for a new measure.

We understand that other practices throughout the country might need a crutch. And we consider this measure really a preliminary step, a crutch measure.

We'd be happy to continue

supporting it if the feeling is that this is necessary. But for us, you know, it's getting close to retiring it in terms of usefulness to New York City. And we'd be happy to work on a new measure such as 0421, or whatever else that we're going to be developing.

MEMBER STOTO: I'd just like to suggest that the one that we should be talking about harmonizing with is 1690, which is the NCQA one, which is really fundamentally the same thing, except for minor differences.

CO-CHAIR JARRIS: So the question here is, would the measures be, would the language of the measures be the same? It would be collected differently, since one's a health plan and one's a clinic. But that gets to what we were trying to say before that we would like to have, the same measure applied differently.

MEMBER STOTO: Yes. But they both have to do with whether or not BMI is documented.

CO-CHAIR JARRIS: Right.

MEMBER STOTO: And that's -- But the NCQA guy may disagree here.

DR. REHM: I'm mindful of Helen's recommendation to kind of -- I've been through a lot of these panels. I kind of live here. Dawn lives here more than I do.

Because it's, you're beginning to conflate the individual measure with the potentiality of where we could go downstream. And quite often those are, you know, sequential conversations.

But, you know, because it's been raised, one difference is that -- You know, this is an interesting -- Because, you know the paradigm is different. New York City is a terrific leader in community health.

And they have, you know, kind of worked that field, and have an EHR specified measure in use with a fairly large population being assessed. Their success at this is quite phenomenal. They're approaching 80

percent.

You look at health plans across the country with a measure that's been in play for three years, I think. And performance is not, you know, it's around 50 percent if you're lucky, depending on Medicaid, commercial and Medicare.

I think that number is probably more reflective of what's actually taking place in the marketplace. So the question is, well my sense of the question. Helen can probably articulate this much better.

Is there a need for endorsement of novel approaches to measurement, that can be adopted by other people who are, in New York City's case, interested in using EHRs in their community to support a measure that appears to be moving quickly? And to be very successful, who knows what's driving the car there?

Is it the decision support system from the EHR? Who knows? But there is a slight difference in our measures. New York

City, because of the EHRness of it all, is just really looking at a height and weight, that then can be calculated by a system.

Again, not with a physician or a clinician interface. And then therefore, in some ways, kind of like the old days of looking at measures, some of the low bar measure of simply documentation. Really, did you document height and weight?

In our measure, you'd have to have actually entered the BMI. You'd have to have calculated it in the medical record. And calculated in the medical record and then put a date next to it. So that you know that someone's performed that function.

DR. BURSTIN: I mean, it's just not that clear cut. I mean, most EHRs just take your height and weight and create a BMI. The last thing I want to do is actually have to pull up the NHLBI calculator, which is how most providers do it, to figure out and then enter it.

DR. REHM: It may be New York City can describe exactly how that takes place. When I read the spec, I interpreted it differently. That there was something else that might be going on.

Because they talked about difficulties getting the BMI calculated depending on the fields. And so it's an interesting lesson on how complicated getting what appear to be simple things done.

CO-CHAIR JARRIS: But if the system -- I mean if you enter a height and weight and the system calculates the BMI, with no clinician, nurse or doctor interface, that is an extremely low bar.

Because then there's no cognitive process going on here. Therefore, the doc may not even know what the BMI is when the patient's in there.

MEMBER KINSINGER: This is Linda. If I could ask a question? Because I'm pretty familiar with how EHRs work in our VHA

healthcare system.

And that is that, while there is a specific field for vital signs, where heights and weights can be entered, they are often entered in the free text note box. And so are not captured.

And we continually reinforce with our providers, put it in the right place, where it's in the calculable and computable, and searchable field. Don't just put it in the note section. And I wonder whether there are, to what extent does that happen? Or is that allowed in the New York City system.

MR. AMIFAR: In the New York City system, we have a section on the vital screen where height and weight need to be entered. They can make it mandatory, so that someone can't leave the screen until they actually do it. That's an option for the particular practice. They can do that or not do that, depending on how they wish.

And within the height and weight

field you cannot put free text in there. It has to be numeric. And there are some out of range boundaries. So if you put a height, for example, eight feet, it will not take it.

You have to put something that's reasonable. Once they've put the height and weight, then the system automatically generates a BMI for the patient.

MEMBER KINSINGER: And so how many practices have made that mandatory as opposed to optional?

MR. AMIFAR: That we don't have data for. It's a setting within the EHR system that they can make it mandatory or not. And I think that's why we have some variability in our numbers.

But there has to be structured data within that field. They can't just type in whatever they want. And we've modified that since, I think 2007, 2008. So that the data in there is valid and interpretable by the system.

CO-CHAIR JARRIS: So I think the question is, is 0023, BMI in adults greater than 18 years of age, 1690, Adult BMI assessment. I mean, the question is, shouldn't there be one measure of this that could get implemented in different ways?

It could be that in New York City, because you put it in the vital stats, it calculates the BMI and comes up that way. It could be that in another setting it's actually a calculation that the provider has to do and date. But it has to be in there.

I mean, it doesn't make sense to me to have two separate measures for these things based on how their implemented. The implementation's secondary to what the measure is we're trying to achieve.

MEMBER STOTO: You're talking about 0023 and 1690, right?

CO-CHAIR JARRIS: Yes.

DR. REHM: Just a comment. So, you know, we've run across this quite a lot.

We have a lot of competing measures out there that we try to harmonize or align. So when steering committees review these, they're reviewing these on the right of criteria you see.

And one of them is where this measure's been tested. And so our locus of interest in this particular measure is health insurance plans across the continuum, commercial, Medicare and Medicaid.

New York City's is on clinicians. A measure specified for clinicians needs to be tested for clinicians, needs to have performance data for clinicians. A measure for health plans needs to be tested for health plans and et cetera, et cetera. Just the corollary.

So, you know, so even if we were to have a harmonized measure, let's say two, three years from now. Wait for the task force guidelines to come out. And we believe that there are maybe two dimensions that are of

interest to clinicians that are exactly the same interest to health plans.

But then because of the nature of health plan accountability, you could very well have a third dimension, or third element. Because of the nature of the beast.

Just like when you're thinking about population health at large, the health of the nation. You may be measuring things quite differently than you would at the level of interest of a clinician.

So in that spirit, you would likely not get what I would call a picture perfect mirror image for clinicians that you might have for a health plan, that you might have for a State Medicaid department, or any other level of accountability that you choose. Accountable care organization perhaps.

So I think that's where this, the idea that you can simply have one without doing the testing, without having the perforance data being generated by the system

of interest, where it gets a little naughty.

Ten, 15 years from now, we may be in a place where the nature of the data and the flow of data moving makes this a moot point. But we're not there. And I think that's the challenge we're facing.

We have one foot in kind of in our, you know -- Maybe it's archaic, but, you know, we believe health plans need to be held accountable in the U.S. And we believe that measures that do that are valuable.

And we also believe that measurers need to hold, you know, providers accountable, as well as communities, you know. But that again needs to be developed over time.

CO-CHAIR JARRIS: Matt.

MEMBER STIEFEL: I mean, that's --

DR. REHM: I'm just looking at -- I'm sorry.

MEMBER STIEFEL: -- an important general philosophical argument. I don't think it applies in this case. BMI assessment is

BMI assessment.

And I don't think that those issues about the difference in specifications of the measure are germane to the -- It may be differences in sample size, and those kinds of other attributes. But not the definition of the measure.

MEMBER STOTO: Just to -- I agree with that. And to continue, I'm looking at 1690. And even though it applies to health plans, it's only of patients who have had an outpatient visit. So that makes it a lot closer to the New York measure.

CO-CHAIR JARRIS: So what the question comes down to is, does NCQA endorse a measure? Or does it endorse measures in the context of the specific means in which they are collected?

Because if that's what's being endorsed, we will have potentially an infinite number of BMI. One for nurse practitioners, noe for PAs, one for MDs, one for specialists,

one for you name it. And that's I think exactly what we're trying to get away from.

DR. BURSTIN: Couldn't agree more. Ideally you would love to have a measure that works at multiple levels. Keep in mind they are often very different data sources. That measures need to be specified differently, for an EHR versus claims, for example.

So we recognize that. And at least for where we are right now, they can be different measures, but they should be fully harmonized. The age group should be the same. The definition of BMI should be the same. The time period should be the same. Is it 12 months or 24 months? Should you have counseling in or out?

I mean, those are all the discussion items we need to go through. So what I would actually recommend we do, just because I think this happens -- It's very easy to get excited about the --

We think harmonization is sexy and

interesting, leap to that. But we actually need to do first I think, is to actually walk through the evaluation of all three measures.

And then walk back and have this discussion of, is there potentially an opportunity to bring some of these closer together, or merge a couple of them? Or figure out what makes the most sense.

But it's hard to do that, because you haven't even talked about the other two measures yet. So let's actually, I would recommend just finish this evaluation and just move on. And then we'll get to the good stuff.

CO-CHAIR JARRIS: Okay. All those in favor say, aye. We then are probably at a point of voting on 0023.

DR. BURSTIN: Yes. I agree.

MS. MUNTHALI: Okay. For importance to measure and report, one --

MEMBER QASEEM: What are we voting then, though? Because --

CO-CHAIR JARRIS: So we're evaluating measure 0023. And we'll go through each of the measures. And then we'll put them in a pile and look at harmonization.

MEMBER QASEEM: But don't you think we need to look at all of them before we vote on them?

DR. BURSTIN: Essentially what you're going to do. I'm sorry. Let me qualify that slightly. When there are competing measures, essentially all you're going to do at the end of this process for each of these three, is indicate whether they have met the criteria for endorsement.

You will then -- You're not going to approve one of them. You will then go through the harmonization exercise. And then you will vote to say which ones you actually approve.

Does that make sense? First, our first stopping point is, we don't really care about harmonization if the measures don't meet

our criteria. Let's get them first, see if they're actually --

CO-CHAIR JARRIS: It's like in a vacuum.

DR. BURSTIN: Exactly.

MEMBER STIEFEL: A slight modification. I wonder if we could have the discussion for all three, and then vote for all three right in a row? I mean, I don't know.

I reviewed the other one, 1690. I don't know if I should be saving my comments, which are obviously equally applicable to this one. Maybe we should just get all of the comments out, and then just vote on them all.

CO-CHAIR JARRIS: Okay. So 1690 we move to then? We'll keep you in suspense, Sam.

MR. AMIFAR: Okay.

DR. REHM: Actually I think the order was to go to CMS. Sepheen Byron is going to join us. She helped with these

measures originally, and has the history.

CO-CHAIR JARRIS: 0023 and 1690, on their face look very similar.

DR. REHM: I just don't --

CO-CHAIR JARRIS: So CMS as a whole --

DR. REHM: Yes. I didn't know if CMS had it's moment in the sun. Had CMS had an opportunity to present it's measure. Yes. That would be helpful since I'm waiting for Sepheen. That's right. That's what we told them.

DR. WILSON: Hi. This is Don Wilson. I'm medical director for Quality Insights. And with me I have Sharon Hibay, who directs our measure and development initiatives.

And we developed this measure under contract with CMS. So we're the actual measure developer, but CMS is the actual measure steward. And we actually have some CMS folks on the line I believe, that may

chime in if need be.

This measure I think was originally developed in 2008, and received endorsement at that time. I think as you folks, and it's been alluded to by Helen and others, it's been --

Since it was developed it's been used now for probably at least three years in the PQRS program for our physicians to voluntarily report on. And most recently now, it's one of the three core measures in HITECH.

So it's one of the measures that all physicians have to report on, in order to be able to achieve their meaningful use incentives. So it's getting a widespread use in that venue.

And I would just like to -- One of the things we were talking about -- It was the New York group for instance. And he was talking about he's with the REC. We're also in REC. So I'm very familiar with what that's doing.

So I would assume in the New York situation, because all these docs are getting the meaningful use, they're having to use this measure anyway. So all the docs that he's talking about are using this in their system, E-Clinical Works.

Being certified for meaningful use is already programmed and set up, and should be doing this automatically. I mean, it should just be happening. So if you're a doc that's going for meaningful use, you are using this measure today, and will continue to do so.

So basically I think everybody's probably familiar with the measure at this point. It is a two part measure. And to answer someone else's question about it being a two part.

Because of the way this measure is reported, it is reported in a granular way. So you can actually get statistics on just whether the BMI was done or not and

documented. There's an entry reported that way.

But they actually get performance for the whole measure, you have to do both pieces. But we do actually collect data on just, for people who just did the BMI, but didn't do the follow up. Or for people who did the BMI, but did the follow up. Or people who did the BMI but didn't do the followup.

So there's several different ways that it's all reported. So we can really sift that granularity out if we're interested in looking at data subdivided or, you know, cut and sliced in that manner.

So basically it's looking at, as you see here, one of the things that our technical advisory panel was very concerned about when we developed this measure early on, was that we not only focus on obese patients.

But really zeroing in on the fact that in the elderly patients, your malnourishment and low BMI is actually perhaps

even more of a problem than the overweight, as far as increasing mortality. So for that reason we really had --

And also there was lots of discussion about whether there really is a different BMI level range for elderly folks. And really looking at the literature. And most of the folks on our panel really felt strongly that there really was a different set of BMI norms for elderly patients.

So that's reflected in our measures. You can see it's broken out by different age categories. So basically our measure is that of looking at the BMI, whether it's documented in the previous six months, either by the provider themselves.

Or they can actually call our measure if they have it, if they're able to show that it was done somewhere else, but they have it documented in their record. And if the BMI was outside of parameters, whether some followup was done.

Again we really wanted to go that extra step to say that they didn't want to just have the BMI documented. They wanted to see if some action was taken based on it. So it has to --

If the BMI is outside of the parameters, a followup plan has to be in the record in order to get full credit for the measure. And if you look at our measure, you can see the components of what constitutes a followup plan. There's things --

A future appointment for followup of the BMI, for edcuation. There's referrals to different kinds of providers, whether a medication was prescribed specifically related to weight, whether dietary supplements were prescribed, excersise counseling, nutrition counseling, et cetera.

So you have to have at least one of those things in your record to show that you did some followip if your BMI is outside of parameters in order to get credit for the

measure. So that's pretty much, I think, a symopsis.

CO-CHAIR STANGE: Why six months?

DR. WILSON: I think that was -- I don't know that we really have a definite strong guideline on that. I think that was just a strong consensus on the part of the panel, that it should be done at least within every six months basis.

CO-CHAIR STANGE: Going down to 18 though, I mean, most 18 and 20 year olds don't come in even annually, let alone six months.

DR. WILSON: Well it's basically if they've -- It's for patients that have been seen during the reporting period. And for the measure, it was originally written for the PQRS program.

So the reporting period is typically a year. So if you're seen in the office within the -- The denominator, you have to have been seen in the office within the past year.

And if you are, then you should have had a BMI within the past six months documented. If you've been seen at any time during a year, during the past, within the six months of that visit you would have had a BMI documented.

CO-CHAIR STANGE: So why not make it a year then?

DR. WILSON: I think that folks felt that being a year would be a bit too long. That you can have a significatn change in your BMI, you know, if you allow it to go longer than six months. That it should be at least assessed --

CO-CHAIR STANGE: So your focus --

DR. WILSON: -- within a six month period of having been seen in the office.

CO-CHAIR STANGE: So you're focusing on the frequent flyers then, right? You're excluding people from the denominator who haven't been in in the last six months?

DR. WILSON: No. It's basically

again, it's a matter of they're seen -- If the patient is seen in the office any time within the year, you want to see whether they've had a BMI within the previous six months. So it wouldn't be a frequent flyer.

CO-CHAIR STANGE: So I came in nine months ago. I had a BMI, but it wasn't in the last six months. So I'm excluded from the denominator? Or I'm --

DR. WILSON: You're in the denominator because you were seen during the year.

CO-CHAIR STANGE: But so then in the numerator then, I actually -- I didn't get a BMI done, because it was done nine months ago?

DR. WILSON: That's correct. You wouldn't be --

CO-CHAIR STANGE: Well that seems wrong, right? Your intent is to look in the last year. And so you're penalizing me if I didn't happen to be in the last six months.

It doesn't --

MEMBER STOTO: It makes sense for part A, but not for Part B.

DR. WILSON: I'm not sure I -- Again, you have to be seen anytime within the year. At the time of that visit, you have to have had a BMI within the previous six months to get credit.

And you only have to have that happen once during the year. So it's to be reported once during the year. So as long as any visit --

If I was seen in the office for any time that year, if at that index visit if I had a BMI within the previous six months, then I comply for the measure.

CO-CHAIR JARRIS: So you came in nine months ago and had a BMI at that visit. Therefore, that BMI nine months ago was within six months of that visit.

CO-CHAIR STANGE: So you're more likely to be compliant if you're a frequent

flyer I guess.

DR. WILSON: Yes.

CO-CHAIR JARRIS: Is there any evidence to support this? That BMI should be done within six months? I've never seen it. But I'm wondering where that came from.

DR. WILSON: Again, as I recall that was a, sort of a consensus of the TEP panel. That was the reason why -- I'm trying to think back at the guidelines. I'm not sure that we have a guideline that ever specifically states how often a BMI should be done.

CO-CHAIR STANGE: You know, maybe --

DR. WILSON: One's 12 months, I think one of the other measures is two years.

MEMBER STOTO: So maybe this should be a harmonization question later. Because one of these measures is six months, one is 12 and one is 24. So that might just be something to think about later on.

CO-CHAIR JARRIS: So when you develop measures, what burden of evidence do you require? You mentioned dietary supplements.

And I'm wondering what evidence was looked at to recommend dietary supplements as a disposition for someone with an elevated or low BMI?

DR. WILSON: I think that there's probably pretty good evidence that dietary supplements help people who are malnourished. I mean, I'm not sure that we have an article that says that, but --

MEMBER KINSINGER: It depends a lot it would seem to me on what the dietary supplements are. I mean, there's a whole huge range of different things thta are all kind of bundled under the broad name of dietary supplements. If you're talking about nutritional drinks like Ensure or something, that's one thing. But if you're talking about --

CO-CHAIR JARRIS: Ephedrine.

MEMBER KINSINGER: Ephedrine. That's another dietary supplement that would not help. Exactly.

CO-CHAIR JARRIS: Don't you have to assume at some point that the physician did the right thing? That the --

DR. WILSON: Right. I mean, I would think that you wouldn't have someone, you know, giving ephedrine to somebody who was under -- You know, again I think the -- And these were really just listed as examples.

In other words, the goal of the measure is to show that some followup is done. It's really not necessarily saying whether it was the right followup.

In all honesty, you just want to have documentation that the physician looked at the BMI and did something about it. He decided to, you know, came up with a treatment plan or care plan.

It's not really assessing whether

there was an appropriate medication prescribed, et cetera. It's just whether a care plan, or a followup plan was done, as a followup to the fact that the BMI was abnormal.

MEMBER KINSINGER: I'm sorry. Can you tell me where the followup requirements are? I'm having a hard time finding it.

DR. WILSON: It's listed under -- If you look under definition, it should be, if you have the measure specs, under numerator. It's in definitions. It says BMI, elderly BMI, calculated BMI, followup plan.

And again, if you look at the language in the measure, it says, followup can include. And it's documentation of a future appointment, you know. Again these are examples of what a followup plan can be.

MS. HIBAY: Just to make sure everyone understands the source document. I think you have two documents, which is the NQF, the submission form of the information we

had to provide. And then there's a supplementary document.

And the first attachment of that supplementary document is, are the measure specifications themselves. And the definitions are there, including what constitutes a followup plan.

In the EHR specification of this measure, when you're doing claims in registry, the way that you report the measure, you're not necessarily saying --

And this is followup specific to upper parameters or parameters above the normals. And this is followup based upon the low normal parameters.

So the definition in the specification is provided as a lump of, this what could constitute, may constitute. But it's not limited to followup.

But in our EHR specification of this measure we have specifically looked at what are those appropriate followup

interventions that would be okay for lower parameters, okay for higher parameters.

Perfect example would be, so if you're in the upper parameters, way upper parameters, you would not, you would potentially refer someone to a bariatric surgeon.

So we have a codification for that measure, which obviously if someone is below parameters, that's not an appropriate followup for them. So there's, you know, there's a teasing out of that information as, you know, as we go forward.

But the information on the specification that you have, it's a lumping of the parameters for upper and lower. Excuse me the followup for upper and lower.

DR. WILSON: But again I want to stress that fact that the measure itself is really just documenting that a BMI was done and that there is some sort of followup.

It's not really assessing whether

it's appropriate followup. Because we got into lots of discussions with that with our TEP. And really decided we weren't --

This was not what this measure was about. Because that creates a whole new set of issues, you know, whether it's appropriate followup.

CO-CHAIR JARRIS: Thanks. Ron and Bobby.

MEMBER BIALEK: I think I recall seeing a reference to the Clinical Preventive Services Task Force in this. And I guess what I'm wondering, and I don't know this off hand. Does the task force itself suggest the followup activities?

I know there's evidence that the followup is effective. And I'm just wondering why the followup piece to this is not just tied to the Preventive Services Task Force recommendations.

MEMBER SPANGLER: Can I comment on that, follow up on that? That was the exact

same question I have. Because the systematic review for the new update, which was published. They have --

One of their questions was, is there direct evidence that primary care setting screening programs for adults, adult obesity or overweight, improve health outcomes. Or result in short term or sustained weight loss or improved physiological measures.

And they didn't identify any trials on screening programs. So the same question that I had for 0023 is why didn't you include screening and counseling in your measure? And they said they willing to do it.

It's the same question I have for 1690. Because there's no evidence that just screening does anything. You do need that component.

DR. WILSON: If I can answer that though. I think what we're saying here is that it's screening. But the fact that it's -

- The physician is now aware of it and actually comes up with a treatment plan, right? So it's actually --

I guess the question you have to assume that treatment for obesity is effective. I mean, so you're wanting to see that the physician recognized it and developed a treatment plan or followup plan for dealing with the obesity. If it's not documented on the chart, you assume that the physician didn't recognized it even.

So the fact that he documents it and then documents that he has a followup plan, that shows that the did something about it and actually created some sort of a treatment plan. And he's going to followup and manage that problem for that patient. It was diagnosed.

CO-CHAIR JARRIS: Bobby.

MEMBER PESTRONK: Set aside the fact that the exclusions are different for each of the BMI measures that we are

considering. And as to whether one set of exclusions is better than another set of exclusions to give us, have the measure give better data.

If 0421, which is looking at both screening and followup were the recommended measure, after we discuss all the measures. And the data were -- And providers or organizations weren't as far along as New York City, or plans that were interested in both screening and followup.

Would the failure to capture the followup information in those places that aren't far along, be problematic? Or would they simply be able to report that first step, which was the first step that New York City took? And be able to decide how well they were doing it, that first step?

DR. BURSTIN: So I think the question actually for QIP is, is it reported out at two separate rates? Such that you would still have the BMI measured rate

separate?

MEMBER PESTRONK: And the reason for asking the question is, if it can be reported out that way, then there's no reason to consider or recommend the other two measures. Because you get more from the third measure, which allows the followup to be counted as well.

DR. WILSON: The way the measure's currently specified, yes. You can approximate out, as I said. If you notice the HCPCS codes that are in the measure.

They actually specify out why you failed. Did you fail just because you didn't document the BMI, but you failed because you didn't do the followup plan. So you can sort out those numbers in the measure.

But again, let me point out, just so we're aware of the context of how this is being implemented in the EMR. Again, for most physicians that are going for meaningful use, they're having to do both things now.

And in order to get your meaningful use, you have to report on this measure, you know, and show that you have performance. So doctors, at least if they're going for meaningful use, they have to report that they're doing both things, you know.

And again, most certified products, again all certified products for meaningful use, which most physicians are implementing, do calculate the BMI.

As long as, I think the one person brought up, the only problem where I could see where it would be not happening is if the physician is not recording it in structured data fields, where the BMI literally could calculate it if they're dictating in and out.

That wouldn't be the case. But as long as they're documenting height and weight they're EMR should calculate a BMI. That should be a given.

MEMBER PESTRONK: Second question. Just to deal with the exclusions then. Why

are there so many more exclusions proposed in 0421 than were proposed in 0023 or 1690?

And does that alter the nature of the validity or reliability, I guess would be the -- Does that alter characteristics of the measure in some way that would be undesirable?

DR. WILSON: In all honesty, when you look at the data as far as how many patients are actually excluded for these exclusions, it's very minimal. It's very rare.

These were brought up again because people on, folks on our TEP, you know, that really had this bias, that they felt it not appropriate for instance to --

You know, if someone has a terminal illness, you know, are you really going to try to talk to that patient about, you know -- And is fair to ding a doc if you've got somebody who's going to --

You know, with the expected life expectancy of two months. Are you really

going to ding him for not talking about their weight, you know, that kind of thing. So they really felt that it was appropriate to have that be an exclusion.

MEMBER PESTRONK: Okay. I was thinking about it in the context of immunizations, where we're finding more and more people say, I just don't want to be immunized.

And everybody says -- Not everybody, but then, you know, the provider says, well okay. So one of the exclusions here is if the patient refuses the measurement.

Is it still okay? Or something to be examined there to see whether the rates of refusal have to do in which the way the process was proposed to the patient?

DR. WILSON: And I think how you deal with that generally is, you know, you can certainly look again to see if the refusals are reported, as far as what they were.

And if you have a doc that's an outlier who has a much higher refusal rate than other folks, that would be a reason you would zero in on that particular provider to say, why does he have such a high refusal rate?

You know, I think my sense of this is, eventually I want -- If we could convince providers that it was okay to be dinged occasionally, I think we could probably be better off. And I'm kind of getting off on my own soapbox here.

But if we could do away with some of these exclusions and just realize that you're going to have an occasional patient that you may get dinged on. But it happens to everybody. Because it really does create such a level of documentation issues in trying to figure out how to define some of these things.

But the problem is that, at least in my experience in working in the quality improvement organization is, it's hard to sell

it to physicians.

You know, when they look at this and they say, wait a minute. And they come up with all these scenarios like that, you know, you're going to ding me if I have somebody who's got cancer.

You're going to tell me I have to tell them they've got to, you know, lose weight. That kind of thing they get -- You can't get their buy in. But if you can tell them, you're excluded, doctor. You don't get dinged for that patient. Then they accept it.

CO-CHAIR JARRIS: Actually I think, Bobby, that's an important point to raise. And I feel really strongly about this, having worked in physician incentive programs.

I thought we were past that issue of docs saying, I can't control the patient. He won't get on the scale. With vaccinations you're exactly right. Or with asthma treatment plans.

We all would say, yes,

occasionally you're going to get somebody, no matter what you do, who's going to refuse to get on the scale, refuse the vaccine. But that's a small number.

And the better doctor will have fewer of those refusers than the poor doctor. So you're going to get measured on it and understanding that once in a while you're going to get dinged. I thought that we were beyond that discussion ten years ago. And I'm sorry to hear we aren't.

Now you're a public agency, so you're, you know, the bait is built into the process. Isn't NCQA beyond that, that whole issue about -- I mean, there's no exclusion for the patient refusing your vaccine rates.

And the second point is, I could care less whether a doctor documents the reason why the person didn't get on a scale. Because that should be such a small number when it's legitimate, that it should just be dismissed and removed as a complication. So

Bob, can you comment?

DR. REHM: Well, you know, we're on record as having very, very few exclusions for the denominator, or exceptions in the numerator for that matter.

But there's clearly attention. Some are more -- The current immunization situation is very problematic, because the number of refusals out there is getting higher and higher, unfortunately.

On the other hand, our rates are fairly stable. So I don't know what's, there's kind of -- The people who do want them, really want them. And the people who choose not to -- So in a funny way it's not affecting the national or regional rates. But underneath that we perceive an issue.

CO-CHAIR JARRIS: So would it help CMS, and if we can do this in the context of this in endorsing a measure, to say, we don't want those two exclusions. Take them out. And then we'll endorse you. Because frankly,

I feel strongly --

MS. HIBAY: Is Dr. Dan Green on the phone still?

CO-CHAIR JARRIS: -- I really dislike those.

MS. HIBAY: Because I understand he called in.

MS. HIBAY: Dr. Dan Green is, I think might be on the extension.

CO-CHAIR JARRIS: Is there a Dr. Dan Green there, operator?

MS. MUNTHALI: Anika?

CO-CHAIR JARRIS: Yes. Dr. Green if you're muted, it would be a help to unmute.

MS. HIBAY: I can't speak for CMS, but I can say that we will present the information. I know that we have been presenting other measures.

And there were requests made by workgroups and by committees to make modifications to the measures. And CMS made every effort to, you know, move those

modifications forward.

So I can say that we can bring that back to CMS if they're not on the line still. I would say that it would most likely be very favorable.

DR. WILSON: So again, if you can just tell me which two was it? Was it the refusal and --

CO-CHAIR JARRIS: Any other reason documented in the medical record by the provider, explaining why BMI measurement was not appropriate. It was Thursday and I didn't feel like asking that day. I documented it in the medical record.

CO-CHAIR STANGE: So I have two process things. But, Bobby, do you have something specifically on this issue before I do that?

MEMBER PESTRONK: I have a question about this, but not the question we've been discussing.

CO-CHAIR STANGE: Let me do the

two quick process ones. One, I've seen a couple of people get up, presumably to go to the bathroom. So wonder if we need a break.

And the other process question is, we're in the process of discussing three separate measures.

And are we getting into enough detail that we're going to have a hard time holding the three in our head to vote on them separately at the end?

And so the question is, should we -- When should we take a break? So raise your hand if you need a quick break. Okay. So a small, but perhaps urgent minority.

And then the other thing, are we getting to a level of detail on these discussions on the three separate measures, that we maybe need to back off to Helen's suggestion and vote on them separately?

DR. BURSTIN: We still haven't talked about he NCQA measure.

CO-CHAIR STANGE: Well that's

exactly the point.

MS. JACOBSON: Kurt, this is Dawn. Could I just have a conversation with you and Paul briefly, during the break?

CO-CHAIR STANGE: Okay.

MS. JACOBSON: As co-chairs.

CO-CHAIR JARRIS: Sure. We'll keep you on speaker phone. No problem.

CO-CHAIR STANGE: My cell, Dawn, is 216-2345-9504.

MS. JACOBSON: Or can we just use the --

MEMBER PESTRONK: So my personal conclusion is, actually the discussion of all three simultaneously has helped me decide which one I think is the best one. And to recommend.

CO-CHAIR STANGE: Are you going to be able to vote separately on all the items for all three? Are people holding that in their heads okay?

CO-CHAIR JARRIS: The other --

Right.

MEMBER SPANGLER: We're not voting on the best one though. We're voting for each of these individually as if the other ones didn't exist first, right?

CO-CHAIR JARRIS: We're not voting to endorse this. First we're voting just it's appropriate. Because one of the things I would think about considering.

And the Legislature used to do this all the time. The legislative committees, they'd say, you know what, the three of you go in that room and talk and come back with one measure.

And I don't know if we can do that also. But I think if they're close enough to say to the three parties, sit down and come back with something that meets the following criteria --

DR. BURSTIN: That's not possible. We haven't even talked about the -- That's the big question. We haven't talked about the

third one yet. So I think we need to talk about the third one. And then we can have that discussion.

CO-CHAIR STANGE: All right. I withdraw my question then.

MEMBER PESTRONK: May I ask then another question? Why, given that this one, 0421, addresses screening and followup. Why wasn't it proposed just to measure BMI and have that be the measure of interest?

So that in the population, whatever that population under consideration is, one sees whether there is a reduction in BMI, where it would be appropriate to have a reduction in BMI?

Why are we focused at the screening and some followup level, rather than at the outcome level, which would be hey, you know, okay, yeas they're screened. Okay, there's followup.

But the real issue is, did BMI drop for those for whom it was appropriate to

drop. Why aren't we measuring that one? Why hasn't anybody proposed that as a measure?

CO-CHAIR JARRIS: That's your next job. Matt.

DR. WILSON: I think that certainly is a very reasonable measure. I mean, this measure was actually though, again developed previously as a measure to really evaluate, as a process measure to evaluate clinicians.

And really trying, the goal was to get clinicians more focused on making sure they are checking BMI and doing something about it.

But I think you're exactly right. I mean, there should be an additional measure to have as an outcome population based measure. It makes perfect sense.

But I think, again I don't think one necessarily negates the other. You know what I mean. I think they each have their own purpose.

Because it's probably very difficult, if you're trying to evaluate an individual practitioner with very small number of BMI patients that have elevated BMI, it may be very difficult to evaluate him statistically if he only has five patients that are like that.

You know, on a physician level it's really very difficult to really distinguish good performers from bad performers versus looking at the whole population. But if you're looking at an ACO or a health plan, then it makes sense to do that.

MEMBER STIEFEL: I reviewed 1690 and was going to save my comment about that until then. But it's exactly that point. I think if this group is going to make a difference, and sort of chart a different path from the standard kinds of NQF endorsed measures, that's the sort of thing that we're going to need to do.

Now we're obviously hamstrung in that we don't have a BMI measure in front of us. But they're sure out there. And it's so close. All of these are about doing the assessment that of course could yield a BMI measure.

It raises lots of different measurement issues about, you know, risk adjustment and that sort of thing. But these are so close.

And, you know, I think Dawn's been making this point all day, is that what we have is a set of health improvement activities, where I'm hoping that our work will, at some point, yield health outcome measures.

And BMI is a nice one. And if they're doing this assessment, they're getting BMI. It's just too bad that we can't just take the next step to reveal it.

MEMBER STOTO: I think I agree in principle with that. But the thing is they're

not getting the BMI on anywhere near a representative sample of the population. So to report the BMIs that come out of the databases that we're talking about here, we're not -- I think could be misleading.

MEMBER STIEFEL: I mean, we do it for everybody. And if the other health plans do it for everybody. And if the public health systems do it for everybody. As I said, the sum of those, or the union of those would be the BMI of the population. So it seems like a good place to start.

MEMBER STOTO: Right. But if you're only doing it on 70 percent or 30 percent of the patients to begin with, which is the situation for people that are not as advanced as Kaiser Permanente, then the BMIs that come out of it is not really meaningful.

MEMBER STIEFEL: That's why I can see for -- If you're looking at clinicians, then the preventative care and screening BMI is a process measure, which they have a small

population, large population, you can look at whether they did those activities.

When you actually get to the health plan level, like an NCQA measure, that's when I'd really want to see there's a statistically significant difference in the BMI of a population being cared for. Because you all of a sudden are aggregating to a large enough level.

And when we used to go to NCQA and have to report on our HEDIS measures, we did have to report statistically significant improvements in certain HEDIS measures. That's where I would look at the leadership of NCQA to come in at that level. Not at the clinician level.

MS. JACOBSON: This is Dawn. And thank you, Matt, for bringing it back around. What I really talk to the co-chairs about is a concern of mine that we're going to be endorsing clinical care system measures, not total population health measures.

And I wanted to clarify what the goal of this committee is. My understanding is we are looking for total population health measures. That those clinical care measures would influence. Along with policy interventions that health departments do, and other things.

So I'm really concerned that none of these use the right data sets that's going to give us a truly, true measure of population health. And why are we endorsing them?

CO-CHAIR STANGE: So, Dawn, I think there's a lot of people that share your frustration. You probably missed the part of this discussion where we said we're actually going back.

These are Phase I clinical measures that we're looking at. And that's the lens with which we're looking at them now. So we shared that. And we'll be discussing that tomorrow as well. I think that's a portion of the discussion tomorrow as well.

So we're just processing.

We're behind schedule. But I think getting all these on the table is what we've decided to do. And I think that we'll save time at the end when we look to harmonizing. So that's the path we're on.

We do need to take a break I think for our transcriptionist's purpose. And I'm using you as a foil for the fact that I could use five minutes myself. So I wonder if at a point we could take five, come back, finish this measure.

And then I think go through the third BMI measure before we get to the child measure. And then do the third measure and discuss our options among these three. Is that, my co-chair, is that reasonable?

CO-CHAIR JARRIS: Yes. Yes, sounds good. How much time are we taking?

CO-CHAIR STANGE: So we'll take a five minute break. Really just a quick break.

MEMBER KINSINGER: Kurt?

CO-CHAIR STANGE: Yes, Dawn?

MEMBER KINSINGER: Hi. This is Linda. I just wanted to let you know that I'm going to have to drop off at 3:30 p.m. And won't be able to rejoin the call.

CO-CHAIR STANGE: Okay.

MEMBER KINSINGER: So I'll be here until then, but then we'll be signing off.

CO-CHAIR STANGE: Okay. Thanks, Linda.

(Whereupon, the above-entitled matter went off the record at 3:10 p.m. and resumed at 3:22 p.m.)

CO-CHAIR STANGE: The five minute break for 13 minutes, we're going to reconvene here, please.

CO-CHAIR JARRIS: Kurt, you're getting us started, right.

CO-CHAIR STANGE: So we're discussing three measures and we have completed our discussion of 0023. I guess the question is are we done with our discussion of

0421 or does anyone have anything they want to ask or say about that before we move on to 1690?

DR. BURSTIN: Since it was raised earlier I just wonder whether anybody wants to further discuss the issues of the evidence for the underlying recommendations for follow-up because that's going to come up as an evidence vote.

CO-CHAIR JARRIS: I think that's a topic that, it sounds like it's worth discussing. It sounds like the CMS process is more of a consensus expert testimony type process, which is evidence but typically a very low level of evidence in terms of the USPA preventive task service.

And that the harder evidence, which lags in time, should come out in the next couple, few months from a new recommendation from the U.S. Preventive Services Task Force.

So we're in this odd period of

time where we don't have the definitive recommendations. And so the question is where do we go here. Do we go with, go ahead.

MEMBER SPANGLER: We don't have the recommendation yet but we do have the systematic evidence review that is guiding that recommendation, which they do cite in their studies of body of evidence.

CO-CHAIR JARRIS: How consistent is the CMS?

MEMBER SPANGLER: They say that behavioral based treatments are safe and effective for weight loss, although they haven't been studied in persons with Class 3 obesity.

Medication may increase. So I think they encompass that as their follow-up plan. But it's screening and some sort of intervention.

The new recommendation is not going to be screening and counseling. It's going to be screening for a management of. So

I think they're trying to be consistent with that with their follow-up plan as a management.

DR. WILSON: And again, if I can just say that this measure is really just to show that there is a follow-up plan. We didn't really want to get down into what is the appropriate follow-up. Because we see that maybe being under a subsequent measure on down the road.

So it's really just these things that are in this measure are really just examples to help a physician understand what they need to have documented in their record to show that they've done something about their elevated BMI.

And we're not really assessing whether it's necessarily the appropriate follow-up, just that they've recognized the, like I think someone else said before, you could have an EMR that's calculating your BMI automatically and you may not have even

looked at it.

But to get credit for this measure you have to show that I've acknowledged it, it's out of parameters, and this is what I intend to do about it.

But we're not really assessing whether it's appropriate or whether it's an evidence based intervention, just the fact that you have an intervention in place.

CO-CHAIR JARRIS: So then developmentally, could that be justified developmentally as the same way we, at one point, were looking at. Was a hemoglobin A1C done in the last year?

Now it may have been 12, but it was done and you got the points. So it that a consistent argument here? Because it would be much better if we said the following evidence based follow-up actions were taken. But that doesn't yet exist.

DR. WILSON: I think the analogy would be, in that case, once the hemoglobin

A1C was done and it was elevated, did you do something about it. Did the physician acknowledge that and say I'm going to do X-Y.

But we're not really necessarily saying that he said he was going to give insulin. Well insulin wasn't appropriate. Maybe he should have done this or that. We're not really evaluating that.

We're just saying that we want to see that at least he did something about it. He acknowledged it and then came up with a plan.

MEMBER STOTO: I would say that just doing the hemoglobin A1C is parallel to the first part of this measure and to 23 and 1690.

So the evidence question then becomes, for Part B of this question, is doing any old thing, is there any evidence that's effective? And, of course, the answer can't be true.

DR. WILSON: So is it premature to

have a Part B?

MEMBER SPANGLER: There are things that are effective and this demonstrates that. And I think what you're asking is, instead of having a follow-up plan, should it be something like evidence based intervention is done. Because the follow-up plan could be I didn't do anything.

MEMBER STOTO: So if the follow-up plan is, I invited them to play checkers everyday, that would not be the kind of thing, I presume, that the Preventive Services Task Force says is appropriate for --

(Off microphone comment)

MEMBER STATO: Right, yes. And so if you have any follow-up plan at all there's no evidence that having any follow-up plan is effective in helping people.

CO-CHAIR JARRIS: I guess where are we in the evolution of this? Because it sounds like maybe 50 or 60 percent of people don't even document the BMI. So the first

step would be just get them to document the thing.

But those who have, could we take a similar process to say, well, if you haven't even recommended follow-up that can't be good.

But if you have recommended follow-up it might be good. So then later, it gets modified to what those follow-ups are. I'm stuck. That's why I'm asking.

MEMBER SPANGLER: I think that, at least in my experience whenever we've dealt with TEPS, when it comes down to having to define a very constrained finite list of what is acceptable follow-up or what is evidence based interventions, that's when you get yourself in a go around circle that you never come out of.

And people are always sometimes very uncomfortable that we've gotten everything on the list that really isn't evidence based intervention, that you've necessarily named everything.

MEMBER STOTO: But I think that what it will come down to is that when we have to vote on the evidence, that just following up in however the physician sees as appropriate, I think we'll probably say is weak evidence.

CO-CHAIR JARRIS: Is it different than counseling?

MEMBER STOTO: That would be follow-up, right?

DR. WILSON: Counseling is follow-up, yes.

MEMBER STOTO: Counseling about what? If they document the high BMI and they say you should lose weight that would count as counseling.

CO-CHAIR JARRIS: Ron.

MEMBER BIALEK: So the U.S. Preventive Services Task Force says screen all adults for obesity and offer intensive counseling and behavioral interventions. And I'm just, again, wondering why that

terminology might not be in the measure?

MEMBER SPANGLER: I think because that is from 2003 and the new one is from what we have. We are in this interim period where we have an up-dated on the evidence. But we don't have the recommendation that follows that evidence yet.

MEMBER BIALEK: Understood. So then the next question would be when it comes to that measure and it speaks to counseling, it speaks to follow-up, why not simply reference, consistent with the U.S. Preventive Services Task Force recommendation?

DR. BURSTIN: I think partly because the measure has to be specified as to what that means, particularly for the HR, which is, I think, an issue.

But one recommendation would be we recognize evidence changes. It changes all the time. So I think, based on the current evidence, and I'd appreciate Jason's read on

this since you've obviously looked at this more carefully, if what is cited there under follow-up management is reasonably appropriate for where we are right now then that's probably okay.

But it would certainly be a requirement that QYP return to NQF with an update to this measure when those USPSTF recommendations come forward if, in fact, there's anything in here that's no longer appropriate follow-up.

The idea of just saying any follow-up management, you may be recommending some things that are inappropriate. And that's not okay.

And that's why, I think, really reconciling with the updated recommendations when they come out in the next few months is really critical.

CO-CHAIR JARRIS: Matt?

MEMBER STIEFEL: Short of that, it creates an interesting circumstance. And the

question is, if we don't have evidence about follow-up, is screening and follow-up better than just screening?

Well, it's at least as good as just screening and is likely to be better. Because some of that follow-up will be useful.

MEMBER SPANGLER: And according to this most recent review, screening doesn't change outcomes by itself. So it is not just as good. It is better.

CO-CHAIR JARRIS: Amir?

MEMBER QASEEM: So looking at this evidence review, I looked at it as well. And I don't think we will ever be able to get into some specific follow-ups. The evidence is just not there.

I was just thinking about some other guidelines. Even if, let's say, if AIC comes higher for Type 2 diabetes you're supposed to start behavioral counseling and weight loss and all that before you go into

the medication management and all that.

But I don't think you can start having a specific measure that says, well, did you do this first. And then did you start out with metformin first before going on with the expensive medications out there.

I think we can keep it just broad, the way it is. And that probably is more logical. Even if that task force does come out with some specific thing that you start off doing this, I think there is going to be so much variation based on the patient, especially when it comes to obesity.

I just don't see that you can have a measure that says that you just do this, this and this.

CO-CHAIR JARRIS: So what Helen was saying, keep it more general now and ask to have it reviewed when the task force guidelines come back. And we'll bring it back if something more specific --

(Off microphone comments)

CO-CHAIR STANGE: So just before we take on the other discussion, we're going to finish this. We're going to then blaze through 1690 because we've surfaced all the issues about this general topic.

Then we're going to vote on the three and then we have two more that we'd like to get done before we end today. So just as we're making our comments, thinking about new topics.

CO-CHAIR JARRIS: Bobby?

MEMBER PESTRONK: So that last series of comments brings me back to the question, is all we should be interested in is whether BMI was measured?

Because what we want to know is from one point of measurement to the next point of measurement was there a change. What happened in between, that might be useful to know and is always subject to what the best evidence is. But the question is whether anything happened or not.

CO-CHAIR JARRIS: If there were a measure that looked at improvement or change in BMI, I would agree with you. You don't care what happens in between the two. But we don't have such a measure.

We don't have a measure of improvement. We have static points in time only. We don't have a change, single point in time. So if it isn't measured then that's a problem.

MEMBER SAMPSEL: So just real quick, Helen, how often in your experience did the USPSTF come out with a public comment, then change their recommendation after the public comment?

So the reason I ask is this. So the current draft recommendation calls for physicians to screen adults for obesity and offer or refer patients with a BMI greater than 30 to intensive multi-component behavioral interventions.

Behavioral interventions include

setting weight loss goals, improving nutrition and increasing physical activity, addressing barriers to change and self-monitoring. So that's what we have as at least what went out for public comment.

MEMBER SPANGLER: Basically it's just more specific than the previous one. It's intensive behavioral therapy.

MEMBER SAMPSEL: Exactly, and basically all of that is also coming out of, if anybody looks at the research coming out of the National Weight Control Registry, that you monitor your food intake by tracking. You quit eating fast food. You get physical activity.

And those are all about maintaining weight. So the evidence is out there. I think this still just goes back to is the CDC measure, CMS measure, sorry, clear enough in that follow-up component.

MEMBER SPANGLER: The thing about that is, as you're aware at CMS, there's a

new coverage benefit around intensive behavioral therapy for obesity that is explicit. It's counseling, the five As, so you could make the measure in tune with that benefit as well.

CO-CHAIR JARRIS: All right, Matt, what's next?

DR. BURSTIN: Based on what you just said is, based on your read of this, is that consistent with this measure?

MEMBER SPANGLER: I think the measure is not as specific as that recommendation. It's just much broader. That recommendation could be included in this measure but --

MEMBER SAMPSEL: I think some of it's there and some of it's there and some of it --

MEMBER SPANGLER: I think NCQA hasn't come out with an updated recommendation. So keeping it a little more general, and saying when that recommendation

comes out we want to look at this again, I think would be appropriate. Because we are jumping the gun. NC, not NC, I keep saying that, I'm sorry.

MEMBER SPANGLER: I think to note about your question to Helen, and Helen probably would agree, there has been much more public comment on recommendations in the past two years than there probably has been in the ten years prior to that.

DR. WILSON: If I can just interject one other thing too. I guess think about, in some of these things that we're talking about, how do we actually measure this.

Like we were saying, does the physician have to use, if you want to go back and audit to say that he did the appropriates report, does he have to use that exact same buzz words that you just said or if he says?

That's where it starts to get very dicey. And now that we're actually getting

into trying to put these things into an EHR, where you have to have a very black and white thing, you have to try to think about how are you going to specify this so that you can really say did the physician, or whoever's being measured, meet the intent of this evidence based intervention, if he calls it.

Do you know what I'm saying? So there's always some judgement and it's really hard to make if you have a very precise thing that's going to fit very neatly into a nice little pigeon hole anyway, if that makes sense.

CO-CHAIR JARRIS: Amir?

MEMBER QASEEM: And again, I think some of it is going to be a judgement call as well. If you're going to look at the task force word to word they're just talking about obesity. They're not talking about overweight, so that's a big difference.

That's why I think we just need to keep some of these things in mind that we're

already little bit deviating away from what the task force is recommending, and underweight as well.

CO-CHAIR JARRIS: Okay, let's move on to 1690.

DR. REHM: I'm going to introduce Sepheen Byron, our Director of Performance Measurement at NCQA, who'll be giving you just a few more details on the measure.

But before that, and I know that we're time constrained, I did want to have a couple of comments based on the conversation.

These measures, both our trials and our adult BMI measure, and our child measure, as you know, has the counseling for physical activity and nutrition as well. But both of these measures were pretty much developed concurrently.

We started work on those in 2004 so we had a long road and I've just been plowing through the literature and all the field tests that were done for the measures,

which is an incredibly large file.

I was trying to answer one question that you had asked, which I'll try to answer later on. But I did want to try to convey to you that about the health plans in the country, commercial Medicaid and Medicare, we have about 118 million members in those health plans who report huge results.

So when we talk about national measures, these measures are implemented and used nationally. I think, as a consequence to that scale, we are inherently conservative.

And the discussion about the report that Jason is referencing, we try to, very much, not do evidence review on our own. We don't think we're terribly gifted at it. We think they're better people at it than us.

And so we do look to guidance from the task force for preventive services, to ACIP for immunization recommendations, to

other entities that we believe are really good at capturing the Gestalt of sometimes discordant information. And because of that, even just talking about what the measure could look like before that guidance has been issued is, in our world, premature. We don't speculate, we just wait.

We can think a lot about the NIH cardiovascular guidelines that everyone is just gasping for. But in truth, we'll just wait. And when they come out we will know what they say.

And we will do our process, which includes measurement advisory panels, technical expert panels that look at only feasibility, and then bring it to our committee on performance measurement with recommendations and sometimes go through that cycle a couple of times before we actually have a measure that's vetted for national use and then, consequent to that, seeking endorsement through NQF.

So I just wanted to have that underpinning understood. Developing measures is not cheap. It's not easy. It's very challenging. We don't do it on dime and we don't do it in a day.

Measures have about got an 18 month from start to finish if everything is good. That's the amount of things, in terms of field testing and then looking at first year evaluations, which we do for all of our measures.

Once they pass that, all of our measures are publicly reported. And many of our measures are used in accountability programs, whether those are public payers or private payers.

So again, conservatism in this case is an advantage because you don't want to hold people accountable for things that are on the flimsy side.

Now, in truth, making a measure from clinical guidelines and making it

adaptable for measurement requires a little bit of finesse. It requires balancing feasibility with scientific evidence, sometimes making some leaps of faith. And so I just wanted to preface our comments with that.

(Off microphone discussion)

MEMBER QASEEM: So you were just mentioning that it's an 18 month process and of course the new evidence comes out. But things don't really change when it comes to a measure. It takes a little while.

I want to, for example, point out to the NCQA performance measures VT prophylaxis, routine prophylaxis, there are enough guidelines that are out there --

DR. REHM: I'm sorry what --

MEMBER QASEEM: Venous thromboembolism prophylaxis. It's sort of side issue. I'm just trying to figure out the process a little bit. They want to know the answer to this.

DR. REHM: That's not our measure.

MEMBER QASEEM: NCQA does have them.

DR. REHM: No.

MEMBER QASEEM: I'm almost positive, Bob, and I'll pull it up.

DR. REHM: It's not a HEDIS measure. Can you tell me where?

MEMBER QASEEM: It's not.

DR. REHM: AMA is the steward of the VT measure.

MEMBER QASEEM: I'll hold the question back, sorry.

DR. REHM: Okay, thank you.

CO-CHAIR JARRIS: Could I ask the folks on the phone if you're not speaking to mute your line? We're hearing a lovely conversation in the background.

DR. REHM: And then just to finish the comment, because of the timing of the 2003 adult measure, pardon me, the adult

guidance, we locked into that paradigm.

And then the USPSTF did a re-review of its child obesity measures. In 2009 they were published. I think it often references 2010.

But again, that gave us an advantage and an ability to adapt our child measure to capture nutritional and physical activity counseling.

So anytime new guidance is issued, we are all over it. And we try to bring those into the measurement and basically evolve the measure.

And we've been doing it for 21 years. Many of our measures started very pedestrian and have become a little bit more sophisticated and more targeted over time.

So you can never guarantee these things but that's essentially the way we do things. So I just wanted to have you understand that before we go into the measure itself.

MS. BYRON: All right. I believe you are doing the adult measure now. So this is the adult BMI assessment measure. It looks like the percentage of adults 18 to 74 years of age with a visit who had their body mass index documented in the year or the year prior.

It's in our HEDIS measurement set. It's also a measure that has been added to the Medicaid adult core set for State reporting.

(Pause)

CO-CHAIR JARRIS: Oh, here we are. Yes, go ahead.

MEMBER STIEFEL: I reviewed this measure, obviously. The comments are very related to the other ones. But just a few points about evidence, we've been talking about that a lot.

And it's not an easy question to answer. You need to think about evidence of what. There's evidence that BMI is associated

with health outcomes of interest, like disease incidents and functional status and life expectancy.

Backing up from that, what's the evidence of these follow-ups on change in BMI? Backing up from that, what's the evidence of assessment of BMI to follow-up to improve BMI to improved outcomes.

And so it gets so we need to follow that causal pathway. So when we're assessing evidence of screening, there's the proximate evidence that something is done about it once there's screening.

But ultimately you need to follow the evidence path all the way to health outcomes. It raises the question about, so you have a measure like this you're evaluating the evidence.

Well, screening doesn't improve life expectancy. But is it important enough to put this building block in place for this epidemic in the country so that you can build

on that building block and follow that causal pathway?

And so it answers a different question about evidence. And it was hard for me to understand what I was evaluating when I was evaluating the evidence. Not providing an answer, I guess just a question.

Then the other comments, I think, are things that already have been discussed. This is a non-geographic visit-based subpopulation measure. But is that a rule out? Not necessarily if you can aggregate these subpopulation metrics.

And then I think we've already had the discussion about assessment versus level. And I believe that the contribution of this committee could be moving more to population health measures as opposed to the clinical measures. And then that would take us to measuring things like BMI.

And disparities, we haven't talked about much today. It was an important

criterion that this group established at the beginning of our discussions, way back in the first meeting that we would evaluate these measures against.

This one performs poorly on that because of the data collection issues associated with this measure. So it hasn't come out much in our discussions about this. But we don't have good information about disparities with this measure.

Last point is, it's noted in the submission, it's due for a 2012 update. We're in the middle of 2012 and there's just this question of sequencing and timing if NQF is endorsing an NCQA measure that's under revision.

DR. REHM: I can answer that question, yes. For instance, we have a suite of four cardiovascular measures waiting for NIH guidelines.

It would be premature, an incredible waste of resources, to pull

together a CV measurement advisory panel to evaluate our cardiovascular measures six weeks, two months, seven months. It's hard to tell when they'll be released, to have them do that work prior.

You wait for those guidelines to be released and then you re-visit the measure. So we would wait until that USPSTF final recommendation is out and work from there.

It's the same for vaccines that are released. It's the same for a variety of things. You wouldn't want to have that panel be operating off of what everyone would admit around the table is old evidence. So I think that's the point.

So my thoughts, Matt, would be that this measure would be re-evaluated in calendar year 2013, which in our terms is next week because of the way we have to plan this stuff out.

But again, we just don't know when that will be released. And again, recall that

we don't like being ahead of the evidence and our CPM votes down measures when we're ahead of the evidence, just flat out. And Helen knows. She's a liaison and she's seen it happen before.

MEMBER STOTO: Three questions, I think the first one's to Helen. Is there a history of NQF endorsing measures, developmental measures? The thought being that we want to start getting people to document BMI so we can build these other measures.

DR. BURSTIN: Those are measures that are typically endorsed for a while and at some point you move beyond them. One example would be VTE prophylaxis.

There was a SCIP measure that said did you order a chemoprophylaxis and then did you administer it. Well, the last couple of years we've said okay, enough. If you're administering who cares if you ordered it. So those are examples.

MEMBER STOTO: Okay, so the answer is yes.

DR. BURSTIN: But over time they get experience, they get high degrees, they move off, yes.

MEMBER STOTO: Second question is for the NCQA people. What do we know about the relationship between BMI being documented and whether the physician actually had it, the result of the calculation, when he or she was seeing the patient? That's the issue that came up with the New York meeting.

DR. REHM: Our measures specify, and I'll let Sepheen weigh in, she's better at this than I. But this measure is specified for both administrative data and then also medical record review, where we actually look at the chart.

And different from some of the other measures, the BMI 28.3 would need to be documented in the chart. So it's not height/weight. It's great if they have

height/weight. It shows how they drove that.

But somewhere a clinician who's entering in the chart is entering the 28.3 in the medical record. There's corollary code sets that can be used as well for that. It means that they were calculated.

Do we know if it happened in absentia of the clinician? No. But in an old paper medical record not a lot of people get their hands on those. It's the attending nurse or the nurse practitioner or the clinical practitioner.

MEMBER STOTO: So we really can't say that just because there's a number in the record that the physician had it at a time when he or she could actually have counseled or taken action.

DR. REHM: I think, then, that would bring into all sorts of things that get documented in a chart. You'd start speculating about is the A1C from the lab, that they read and is entered in by hand. But

I guess I really don't know the answer to your question. We do take it on the merits that a clinician saw that number and acted accordingly. Remember, we're having 35, 40 percent rates of just documenting it.

So those are fairly low. Now, they're improving and you'll see in the data we've presented they're improving quite dramatically, especially Medicaid plans where there's a lot of emphasis. So that's the good news. But it is a glass half full.

CO-CHAIR JARRIS: Kurt and then Matt.

CO-CHAIR STANGE: I say this with a lot of humility. I'm less than awed by the logic of consistency of some of our arguments here, my own thinking on this particularly. The first two measures that we discussed this morning, the population measures on being foreign born and the year you came to the U.S., we rejected the building block argument.

And yet here we have a building

block argument, actually bolstered by the fact that Jason just says that the recent evidence review says there actually is no evidence of effective outcomes from measuring BMI.

MEMBER STOTO: I don't think that's quite an appropriate analogy. Because for that foreign born one, the issue at best would be a strata. The documenting the BMI is the first step towards actually taking action. So I don't think it's analogous. But my question really is how close does having it be in the record correspond to actually the physician being able to take action with it. And I guess it varies, is the answer.

CO-CHAIR STANGE: That was actually not my statement. That was my provocative opening statement. But if you want to follow-up on --

MEMBER STIEFEL: What you just made is an argument in support of 421 because 421 requires the clinician to be aware of the BMI.

MEMBER STOTO: And then my third

question was going to be, again to NCQA, is what you guys think about 421 and the discussion that we had about it?

DR. REHM: Look, measure development is hard. I've already talked to that. I think anyone who's in this field should be congratulated and supported and nurtured.

We don't sit around and debate the other measures. I think it is important that we did mention that there was "competing measure" in our submission. Because we did take a turn around the block.

And you just may want to see how that's handled in other cases. But we're highly aware of this. We begin, in many ways, harmonizing before the fact because Helen has us so well maze-trained that to not do so is to make life difficult after the fact.

So we've done this in many settings, in many different clinical areas, and we will continue to do that. It's a big

sandbox. We all want to play in it. And we've worked collegially with all sorts of folks and measure developers. And we look forward to the opportunity to do so.

MEMBER STOTO: I wasn't asking you to criticize or to praise that measure, but just about the idea that you need the second step for it to have evidence that it produces good outcomes.

MS. BYRON: I would say that you should remember that this is a health plan level measure and theirs is a clinician level measure. So it's about what is the accountability of the health plan versus the accountability of the physician and what are you measuring.

So for this one we wanted to create a measure that could be pulled from administrative claims or from medical records. It's a hybrid measure.

And we saw it as a building block of at least look to see that your members' MBI

is documented. And looking at the rates being in 25 to 30 percent, especially for Medicaid plans, it's clear that first step still needs some improvement.

And I think measures across different levels can be aligned without being identical. And the considerations that come into play are who's the accountable entity and whether the data sources that are available, what are the processes, the work flows. I'm sure I don't need to remind this panel of all those issues.

DR. REHM: Yes, if I maybe just cite an example, not to compare this event. We talked before about the measurement the period.

And it's two years for the NCQA measure and why that would be caught by a health plan. Because whether we want to admit it or not we have an overuse problem in America.

And we don't want to have a measure

that essentially is trying to drive people into non-evidence based frequency of visits for periodic health maintenance or health exams when, in fact, there's no evidence that would support that.

So for the adult measure, we have a fairly broad period, which is two years. Do it any longer you begin to drop out the people.

Do it any shorter you're basically saying, boy, if I want to move my measure up, pretend I'm being paid for performance, then really what I need to do is I need to bring in adults all the time and just crank out that BMI documentation.

Well, the measure intent wasn't intended to do that. So this is where we have to balance these things. And that viewpoint may not be caught at the physician level, at a clinician level.

As a clinician, you may want to say I really do want to see everybody on my panel

once a year. But from a health plan payer and an evidence based perspective, you may say that's probably not appropriate.

So that's an example maybe where we could see, depending on the level of accountability, your measure specifications may shift. Or there'll at least be different tensions out there.

CO-CHAIR JARRIS: Amir and then Jason.

MEMBER QASEEM: I'm just trying to simplify it in my mind. You guys are very bright and smart and maybe you can help me out.

So you screen for disease, only those disease or whatever you're screening for when the treatment is available, right? That's the basic rule. And whenever you're screening you want to make sure the treatment is provided.

So when we were talking about even HIV this morning, you're screening so a timely

treatment can be provided. In obesity case, I was looking at the task force recommendation. They're recommending to screen so you can provide certain intervention.

There is no benefit of just screening for something if you're not going to provide that intervention. You're not going to have a measure for any of the preventive services if you're not going to be also talking about the intervention.

So I feel that both of them need to be together. I don't think we can have a, and again, in my mind I'm really simplifying. I understand the health plan level and provider level and all that.

But the point is to calculate BMI because through just one index I really am having a hard time grasping all these various levels. It's a BMI and it is a measure of obesity. And the point is to get the population level obesity down.

It's as simple as that. And I

think we probably do need to have both together rather than separating them out. And am I really over-simplifying this or is this how you're --

MEMBER SPANGLER: I was going to say the same thing. I agree with you that rates are low at 30 percent. But there's no evidence that if we get them up to 100 percent that changes anything. And there's no evidence that shows that it does it. So there's no trials or anything.

MS. BYRON: And I would agree with what you're saying. I think if you think about measurement, you have a disease or you have a condition.

You have obesity. You have a logic model, which was well laid out about how you document your BMI, you look at the BMI, you assess it. You discuss a follow-up plan.

All along the logic model I would say where do you want to measure. And so we've chosen one point along the logic model

to measure. I don't think it means that we are devaluing any other point along the logic model.

I think what we're saying is here's where we want to put our measurement efforts. And, when it gets to the point where all of those rates are high, they look good, then we might want to turn our measurement efforts towards measuring a follow-up plan or measuring treatment or measuring, really ideally, the outcome. Did your memberships' BMI rates fall, period?

MEMBER QASEEM: So if you're looking at one point in time, you do agree, that's exactly the point I'm making. It's just one point in time you're measuring, which is giving you incomplete answer. You're not really getting to what you're supposed to be doing. You're saying that this is Part A of this, whatever measure, and we need a Part B as well to make it a complete measure. So you see my point?

MS. BYRON: I do see your point. I think that in an era of scarce resources and deciding where you want to put your measurement efforts, we've chosen to put it into the documentation at this time.

But I definitely agree that all along the point of addressing obesity, treating obesity, all of those are important. But this measure looks to see that there was documentation.

It's a health plan level measure. It's looking at a health plan population. And this is where we believe that the point of care can be improved.

CO-CHAIR JARRIS: Sue?

MEMBER PICKENS: So I just want to talk at this point about the importance of harmonization. Because we're a large system of multiple clinics that have many health plan members, yet we're also very involved in meaningful use.

And to have to go about trying to

meet both measures is difficult. So the harmonization is really, really key here. And we're also a medical home, which is an NCQA thing, so the follow-up piece is really also important.

CO-CHAIR JARRIS: If I can jump in, I'm trying to de-complexify this. Is that the word? And I appreciate your comments to do that.

To me I see two approaches we could take. We could say, well, since only 40 or 50 percent of the time is the BMI documented, let's take an incremental approach and say, okay, we'll go with that measure, understanding fully that when that improves and there are other recommendations we can add on Part B.

So that's one approach. The second approach is to say what you said, Amir, why would you do a test if you don't do anything with it. That's a principle in medicine. You don't do it if you're not going to take action

on it.

And therefore, why don't we take that two-part measurement, similar to what CMS did to say Part A, it's been documented. Part B, there's some awareness and action taken. So we could look at that other one. We're saying, okay, the PMI's done.

The second thing, of if you're looking for there to be some action taken upon it, a referral or whatever, is it simply an indicator of was the clinician aware of it? Did it go into a thought process?

And then do something, And we could take that approach understanding that needs to come back as soon as the task force comes back with some more specific recommendations for a more specific look at what action was taken.

So I can see either of those things being justified at this point. But I think we've got to start narrowing this conversation down. Because it's after 4:00 already.

MEMBER STATO: Can I speak on that point? And I think that's the right way to frame it. And I think that the justification for 23 and 1690 and the first part of 421 is the developmental idea, that we have to start doing it now so we can eventually add on later.

Adding on the second part of 421, really only says that the physician was aware of it. Because just follow-up in any old way, to put it bluntly, is not consistent with what the task force is recommending.

So there's really no evidence that follow-up that is as general as this will lead to outcomes. So the real benefit of that second part of 421 is that you're sure the physician is aware.

DR. BURSTIN: It's also just not physician. I just want to point that out. This measure is specifically all clinicians, nurses, PTs, everyone.

CO-CHAIR JARRIS: I think, Matt,

you were next and then Kurt and then back to Sue.

MEMBER STIEFEL: I think that the burden of evidence for this ought to be more modest than what we're laying out in terms of the relationship between screening and obesity.

I think that the burden of evidence for this is does increased screening lead to increased follow-up. And we don't have a lot of evidence about that but there's some face validity.

And there are so many measures that we're subjected to that are exactly the same, measuring hemoglobin A1C or lipid levels or blood pressure. All of those don't do anything to outcomes in and of themselves. But the causal pathway is articulated well enough to recognize that first step is a fundamental building block. And I think that's the case with this one.

CO-CHAIR JARRIS: We should start

narrowing down here.

CO-CHAIR STANGE: A question for Sepheen. Is there anything about the child measure that can help us to understand this measure?

So you made a different decision on that measure to include counseling. Is there anything in the evidence for children that can help us understand your decision in this measure?

MS. BYRON: For children there are a couple of differences. We're actually requiring a BMI percentile and we are requiring the counseling for physical activity and counseling for nutrition.

I believe we felt that, at the child level, that was really important. The task force does recommend screening children and offering, or referring for comprehensive, intensive behavioral interventions to promote improvement in weight status.

And we felt that it was more

appropriate for children at the time that we were developing the measure.

MEMBER STATO: It does or does not say that for adults?

MEMBER SPANGLER: It does.

MEMBER STATO: It does, so it's the same burden of evidence.

MEMBER STATO: It's the same but they felt it was more --

MS. BYRON: And the opportunity of a --

MEMBER STATO: -- the same evidence base is there for adults that could be included here already.

MS. BYRON: Opportunity for improvement for children was seen as higher.

MEMBER BIALEK: I have another question related to 1690. You said that this is a health plan related measure.

And while I am understanding the arguments about not measuring the screening part here, that you're putting your emphasis

on the data for the BMI, my question is at the plan level why does your measure not necessarily address how the plan uses the BMI data that the plan receives?

Because, presumably, you're collecting the data for a purpose. And then if you're not having the clinician do the follow-up, what's the follow-up for the plan?

MS. BYRON: Well, I think we would go back to the issue of what do you want to measure versus what do you expect to be done. We have heard from health plans and we have an annual meeting where health plans come together and share best practices.

And there have been several examples of how health plans have looked at BMI and used this measure to understand where they needed to do quality improvement efforts and where they needed to do some drill down analyses to understand where they could improve.

And so we have seen where plans

have used this measure to do quality improvement. We're not necessarily measuring it but we are hoping that this measure provides information to kick off some quality improvement.

And I think that's the case with a lot of HEDIS measures. For cervical cancer screening we measure just the screening. We're hoping that when people have really low screening rates that they go back and that they see that their population needs improvement and they organize their quality improvement efforts around that.

Some measures get pulled into health plan accreditation as well and another way that we try to link the measures to action.

CO-CHAIR JARRIS: Sarah, then Bobby.

MS. SAMPSEL: So this is just a general statement that has nothing to do with the NCQA measures. But when you look at

guidelines and obesity recommendations with adults the overall general guidelines, whether they're from AAP or Internal Med or others, typically follow USPSTF.

However, for kids there's AAP guidelines and Bright Futures. And in Bright Futures they clearly recommend, as a standard of care, nutrition counseling and physical activity counseling. So that may be why NCQA chose to include those in their measure.

CO-CHAIR JARRIS: Bobby?

MEMBER PESTRONK: How would the measure be written? What would it say? I'll approach this in a different way. It's either a point in time measure or it's a measure of what happens from one point to another.

And it seems to me that these measures are focused at a point in time to see whether something was done. And what we're really looking for is something that will let us know whether things changed over time.

When I asked earlier why don't we

just have as the measure the BMI, because if we have the measure of the BMI then we will know whether things change from one point to the other with respect to the BMI, which is ultimately what we're interested in. So if that's what we're ultimately interested in, how would the measure be written? What the language be so that it would be the BMI and not whether the BMI was measured? Is that making any sense?

MS. SAMPSEL: There would be a number of ways to do that. One would be almost the method used by PQRS where you could report a G-code. So you could actually do it administratively where a G-code would equal a certain level of BMI.

So the measure would be out of, well, just go with adults, out of adults 18 years and older. Report by segment, which would be the numerators where someone falls in the BMI category. And so you could do it categorically.

Or you could do it as an actual statement. The problem is, then, that's a specific patient-level measure. And you still have to get to the aggregate level, which is why you might want to do categories. So the measure would either be percentage of members that fall within X BMI category, overweight, obese, whatever, and do it categorically.

Or you would definitely have to get down to patient-level data, which most people don't do unless you're doing a survey.

CO-CHAIR JARRIS: So, Elisa, the UNS operator let someone else on. After this comment can we ask someone to call the question, formulate a question and call the question here.

MS. MUNTHALI: Anika, Anika?

OPERATOR: Yes, hi.

MS. MUNTHALI: Is Mary Barton on the line? And if she is could you please open up her line? She's been trying to reach us.

OPERATOR: And her line is open.

MS. MUNTHALI: Okay, thank you, Mary?

MS. BARTON: Thank you, this is Mary Barton, Vice President for NCQA's Performance Measurement Department. And I think that this conversation has been highlighting the fact that clinical care of obesity is currently, notwithstanding the upcoming U.S. Preventative Services Task Force recommendations, still in maybe a pre-penicillin phase, we'll call it.

We have to get a lot better at figuring out how to help people make desired behavior changes so that they can change their health trajectory with regard to their weight.

And the fact that we would like for these things to happen, we would like to have better tools than we have, but I don't think that necessarily means that we're in the right place to tell people precisely what they should be offering to every patient, given the sparseness of the research on what works for

what people and in what setting. I think that's all. I'll stop with that.

CO-CHAIR JARRIS: Okay, so who speaks next with our recommendation?

CO-CHAIR STANGE: I'd like to support Helen's point of voting on the measures. The two outstanding questions I still hear on the table are whether it's sufficient to do the BMI measurement or whether that needs to be linked with counseling.

And then the other one that Sue brought up is harmonization across the measures. And what I think I hear Helen saying is the way to create a forum for the measure developers to be able to work on those two issues is for us to vote on the measures.

DR. BURSTIN: And only those that are appropriate for endorsement, that meet the criteria, will then go on to the next discussion about harmonization.

MS. MUNTHALI: So the first measure

will be the New York City measure. This is 0023. So for importance to measure and report, impact IA, the rating scale high to insufficient evidence, one to four.

And we have 20 seconds. And remember to point your clicker to Kristin's monitor. Are we nine now? Okay. And, Linda, oh, she signed off. So high five, moderate three, and one low.

CO-CHAIR JARRIS: I didn't get in. Put me into moderate.

MS. MUNTHALI: So there's a correction, high five, moderate four, and one low.

CO-CHAIR JARRIS: I don't know why we had to high five, but we did it.

(Laughter)

MS. MUNTHALI: So 1B, performance gap, same rating scale, 20 seconds. We have all ten. So eight have voted for high and two moderate.

1C, evidence, this is yes one, two

no, three insufficient evidence. Four yes, two no, four insufficient evidence.

MEMBER STATO: What's the difference between no and insufficient in this case? So the question is whether yes is, more than half of them are yeses.

MS. NISHIMI: For purposes of tallying it doesn't. But obviously some people feel it's a flat out no. And others just think there's not enough. I think voting there would be a distinction but when we collapse them into the roll up category of pass, no pass then it fails.

MEMBER QASEEM: Because this is something that's been debated enough in the guideline business and performance measurement world I think it's important to look at it.

Because if you're categorizing it together just like what you just said. That no and insufficient is collapsed together and maybe they should be collapsed together.

MEMBER STATO: So this distinction

between I and C in the task force recommendations, I understand what that means.

MEMBER QASEEM: Yes, side issue.

CO-CHAIR STANGE: Before we go on I just want to point out that we're voting not to recommend a previously endorsed measure. And so what's the logic for that. It's logical if we actually take into account what Jason said, that this evidence report now says that there's no evidence of benefit from just measuring this. So that would be a logic for this.

And if that's the logic we're voting with then that's fine. If not, then we might want to talk about this.

DR. BURSTIN: And also our criteria changed and became significantly more stringent than they were the last time the measure was endorsed.

MEMBER QASEEM: Do we need to vote on it, that we agree with Jason's, whatever we

call it?

CO-CHAIR STANGE: Just saying that would be a rationale for why we would not endorse a previously endorsed measure.

MS. MUNTHALI: So the next measure is 0421. This is also a BMI measure from CMS, also endorsed. Importance to measure and report, IA, impact, leaning scale, one high, two moderate, three low, four insufficient evidence. I think one person, good, we have everyone. Six highs, four moderates.

1B, performance gap, same rating scale, 20 seconds. We have everyone. Eight high, two low. And 1C, evidence, rating scale yes, no or insufficient evidence. Eight yes, one no, one insufficient evidence. So we'll move on.

This is scientific acceptability, reliability, 2A, high to insufficient evidence. We didn't get all the votes unless somebody abstained. I don't think so.

We actually do need everyone. If

you abstained then you abstained. But because we have ten as our quorum, if we can do that again. One more, okay, we got it. Three high, six moderate, and one low.

Validity, 2B, same rating scale, three more, two more. Three high, six moderate and one low. Feasibility, same rating scale. Three high and seven moderate. And feasibility, same rating scale. Five high, four moderate, and one low.

And so now we're voting on the suitability for endorsement. And this is a yes and no, yes or no, sorry. Ten yeses.

So the next measure is 1690, an adult BMI assessment measure from NCQA. It's a newly submitted measure.

(Off microphone discussion)

MS. MUNTHALI: For 1690, importance to measure and report, 1A, impact, rating scale high to insufficient evidence. I think the timer has started. Okay, we got it. Five high, two moderate, and two low, and one

abstention from all of the voting for 1690.

Okay, 1B, performance gap, same rating scale. Six high, two moderate, one insufficient evidence. 1C, evidence, yes one, two no, three insufficient evidence. Two yes, two no, and five insufficient evidence. So we stop here.

CO-CHAIR JARRIS: Matt?

MEMBER STIEFEL: Just a comment about the voting about evidence. It's related to a comment I made earlier about if it's evidence from all the way from the assessment to the outcome, and that's how the evidence is generally written, I think that's reflected in this voting.

If it's evidence between assessment in doing something, that's a more modest basis for acceptance. But it's not in our frame. So it's tricky to do this evaluation.

Evidence of what is the question. And I think that we don't have a lot of clarity. If the evidence is about all the way

to the outcome, boy, I can think of all kinds of measures that don't meet that rigid test.

MEMBER STATO: I would agree with that and point out that, even though this is still in our Phase 1 work, when we deal with these population health measures that's going to be an even bigger issue, sorting out what's the appropriate level of evidence.

CO-CHAIR JARRIS: Move on to children. And this is NCQA again, if you'd like to speak to it. And then who in the committee reviewed this one? Jason, okay.

MS. BYRON: All right, this is weight assessment and counseling for nutrition and physical activity for children and adolescents. It looks at a BMI percentile and at counseling nutrition and physical activity ages 3 to 17 years old. There's a look back period that means you're capturing 2 year olds in the measure. It is included in the Medicaid childrens' initial core set of measures and also part of meaningful use.

MEMBER SPANGLER: So, yes, I looked at this measure. I think all my comments are in the memo. There are a couple issues that I have. One is just the explanation for why the age of 3 when the different guidelines had differing ages.

The task force has 6, AAP and Bright Futures has 2. So just wanted to know the explanation for that. And the developer did provide that. They went with the AAP and then added a year. But there's no real evidence for what the lower age limit should be.

One question, which I think was addressed already with the previous discussion, was should there be two separate measures of this or should all three, I see this as three components.

One is a screening, then there's counseling for nutrition and counseling for physical activity. And should it be that all three of those are into one measure or should

we have separate measures.

And they could be two separate measures, one with screening and counseling and nutrition, one is screening, counseling and physical activity.

I just brought that up as a question. I'm not advocating for that but I don't know if there should be discussion about that.

And then the last issue I had, this is just measuring a BMI or a percentile. And then my understanding of the evidence, and someone correct me if I'm wrong, is the evidence isn't that great for counseling if you have a normal BMI.

So for a lot of kids, my wife's a pediatrician, they still talk to them about what they should eat and whether they should do physical activity.

But my understanding is there's no evidence that changes anything, that it maintains current BMI or not, that the

counseling actually, even in children and adolescents, is similar to adults, that for those with an abnormal BMI, whether it's high or low, is where that works.

So I didn't know if there needs to be a stipulation there or the measure just has to be adjusted for abnormal BMI or not.

CO-CHAIR JARRIS: On the other hand, it's hard to imagine any downside to giving anticipatory guidance to somebody with a normal BMI about eating and being physically active. So since there's no downside I wouldn't worry about it being done.

Any other questions or comments or anything unique or different about this measure that we haven't beaten to death in prior measures? Robert?

MEMBER QASEEM: Can we get a response from NCQA regarding the age please?

CO-CHAIR JARRIS: Yes.

MS. BYRON: So Bright Futures does recommend it for age 2 and our measure

captures age 2. It's written as age 3 by December 31st of the measurement year. So you actually do capture the 2 year olds.

So if you turn 3 in December you were 2 for most of the measurement year. So that's our way of capturing the appropriate age range. And most HEDIS measures are structured that way.

MEMBER SPANGLER: So when you have an age range for any of your measures it means turning that lower limit age by the last day of the year?

MS. BYRON: Yes. So for breast cancer screening you've turned 42 by the end of the measurement year. But it's a two year measure with a two year look back so you're capturing the 40 year olds.

CO-CHAIR JARRIS: Okay, Bobby?

MEMBER PESTRONK: This one says that the data source are paper medical records. But I assume that something happens to those records to make them electronic so

that things can be counted.

MS. BYRON: No.

MEMBER PESTRONK: No?

MS. BYRON: Plans to do a medical record review. It's a hybrid measure and they do pull a sample and look at medical records.

CO-CHAIR JARRIS: I guess it's the word paper Bobby's commenting on. Because you would also look at an electronic health record also, right?

MS. BYRON: Yes.

MEMBER PESTRONK: And then the denominator statement says the primary care physician or ob/gyn, there's a specific listing. This is controversial, I guess, in the medical field. Who is a primary care physician and who isn't. So is there some good understanding of who that is?

MS. BYRON: Well, the measure provides codes that basically defines, all HEDIS measures that are relevant for primary care apply to primary care physicians. And we

include ob/gyns in that as well. Because in some states your ob/gyn can be a primary care physician.

MEMBER PESTRONK: Is there any specialty group that's excluded because they're not considered a primary care physician? And would someone who saw a nurse practitioner, for example, be considered a primary care physician?

DR. REHM: I think in our guidelines we characterize primary care practitioners. The PCP stands for many things.

In this case it's primary care practitioner. My recollection is, and I'm pretty sure of it, that this includes physician assistants and nurse practitioners.

MS. BYRON: Yes, so my apologies for using that word, physician.

DR. REHM: But you're right. We're not measuring the happenstance of a cardiologist who happens to have a young

patient roll through, that they did the BMI. We're trying to capture the primary care physician, that's the intent.

DR. BURSTIN: And I'm not sure if Sepheen mentioned it but this measure has also been retooled for EHRs already, as part of the meaningful use program.

So we say medical record but, hopefully, since there are specs there, that EHR is where it will, hopefully, be going.

CO-CHAIR JARRIS: Okay, any other questions or comments? Anyone on the phone from the committee? No, there's not anymore. Okay, got your clickers ready?

MS. MUNTHALI: Okay, importance to measure and report, IA, impact, high to insufficient evidence, one to four. Oh, Sue.

CO-CHAIR JARRIS: Oh, we need Sue.

MS. MUNTHALI: We need Sue.

(Off microphone discussion)

MS. MUNTHALI: Okay, again, 1A, impact. Eight high, one moderate, and one

abstention for this and for all the other voting on this measure.

1B, performance gap, same rating scale. Seven high, two moderate. 1C, evidence, yes one, two no, three insufficient evidence. Eight yes and one no.

Scientific acceptability of the measure properties, 2A, reliability, high to insufficient evidence. Six high and three moderate. 2B, validity, same rating scale. Four high, three moderate, and two low. Usability, same rating scale. Five high and four moderate. Feasibility, same rating scale. Three high and six moderate. Suitability for endorsement, yes or no. Eight yes and one no. The measure passes.

CO-CHAIR JARRIS: Well, very good. So we have gotten through this part of the day and the process with these measures. I think it will be interesting for us to think about the outcomes here and why we thought they came out the way they came out, which has

implications for future looking when we look at different measures.

We, at this point, should go to the public comment period. And then we can put the background information for Day 2 off to the morning, delay that a little bit, except for the one thing Kurt wants to, excuse me --

(Off microphone discussion)

CO-CHAIR JARRIS: Operator, are there any public comments or anyone in the room? Okay, we'll take that as a no. Kurt?

CO-CHAIR STANGE: I just want to offer people the opportunity for any additional questions for us to sleep on to frame the discussion for tomorrow.

We have two pages up on the wall here and we have a set in the agenda. And I guess that includes, after you've slept on it, when you come in for breakfast if you want to add anything to the list feel free to do so.

CO-CHAIR JARRIS: Okay, any other comments today?

MEMBER PESTRONK: The third bullet on the easel, isn't the question about it's actually a question about what are they smoking rather than is somebody smoking something here which is making this difficult to do, right?

(Laughter)

CO-CHAIR JARRIS: Who's the they? Definitely been ambiguous. Okay, so logistics please.

MS. MUNTHALI: Logistics, since we've run a little late today, we will do the background information for tomorrow's discussion first thing in the morning.

We are having, after that, a facilitated discussion. Kurt, Helen, Robyn, Paul and I are going to talk to the facilitator right after this. And that's why we needed to cut our day a little short.

And then we also offered an opportunity for folks to have dinner together. If you're interested in that just let us know.

We have our colleague who can help to arrange a place or we can offer suggestions for you.

We just wanted to thank you for joining us today and see you tomorrow. Oh yes, same time, 8:00 a.m., breakfast. And we'll start at 8:30.

CO-CHAIR STANGE: And tomorrow we really will start at 8:30 because there were no plane and other delays.

MS. CHANDLER: One other thing, if you could just leave your remotes. Thank you.

(Whereupon, the above-entitled matter went off the record at 4:46 p.m.)

C

H

R

U

C E R T I F I C A T E

This is to certify that the foregoing transcript

In the matter of: Population Health
Steering Committee

Before: NQF

Date: 05-30-12

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