NATIONAL QUALITY FORUM

+ + + + +

CHILD HEALTH QUALITY MEASURES STEERING COMMITTEE

+ + + + +

TUESDAY
NOVEMBER 9, 2010

+ + + + +

The Steering Committee met at the National Quality Forum, Suite 600 North, 601 13th Street, N.W., Washington, D.C., at 8:00 a.m., Thomas McInerny and Marina Weiss, Co-Chairs, presiding.

PRESENT:

THOMAS McINERNY, MD, Co-Chair
MARINA WEISS, PhD, Co-Chair
MARTHA BERGREN, RN, DNS, NCSN, National Association of School Nurses
SARAH BROWN, MSPH, The National Campaign to Prevent Teen and Unplanned Pregnancy
CARROLL CARLSON, RN, BSN, Group Health Cooperative of Eau Claire
ALEX CHEN, MD, MS, Keck School of Medicine
DAVID CLARKE, MD, The Children's Hospital
JAMES GLAUBER, MD, MPH, Neighborhood Health Plan
MARGARITA HURTADO, PhD, MHS, American Institutes for Research
KATHY JENKINS, MD, MPH, Children's Hospital Boston
PHILLIP KIBORT, MD, MBA, Children's Hospitals and Clinics of Minnesota
ALLAN LIEBERTHAL, MD, FAAP, Southern California Permanente Medical Group
MARLENE MILLER, MD, Msc, Johns Hopkins Health System

DONNA PERSAUD, MD, Parkland Health and Hospital System
JAMES QUIRK, MD, PhD, Stony Brook University Medical Center
GOUTHAM RAO, MD, University of Pittsburgh School of Medicine
ELLEN SCHWALENSTOCKER, PhD, MBA, National Association of Children's Hospitals and Related Institutions*
BONNIE ZIMA, MD, MPH, UCLA Dept of Psychiatry

NQF STAFF:

HEIDI BOSSLEY, MSN, MBA
HELEN BURSTIN, MD, MPH
EMMA NOCHOMOVITZ, MPH
EUGENE CUNNINGHAM
SUZANNE THEBERGE, MPH
REVA WINKLER, MD, MPH

ALSO PRESENT:

SEPHEEN BYRON, MHS, National Committee for Quality Assurance
KERRY CHRISTIANSON, American Medical Association*
MARY McINTYRE, MD, Alabama Medicaid Agency*
KAREN PIERCE, MD, AMA Physician Consortium for Performance Improvement*
JANELLE PLUMMER, MD, American Academy of Pediatrics*
COLLEEN REULAND, MS, Child and Adolescent Health Measurement Initiative*
SARAH SCHOLLE, MPH, DrPH, National Committee for Quality Assurance
SAMANTHA TIERNEY, MPH, AMA Physician Consortium for Performance Improvement*

*Present via telephone

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

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

8:08 a.m.

CO-CHAIR MCINERNY: Welcome to the second day, and congratulations. We actually have reviewed 19 measures, and you have, I think at your place you have the list -- oh, we just have the list. We'll give you one.

But we did review 19 measures, with 21 to go, so we're almost halfway there. And I think the pace definitely picked up a little bit in the afternoon, and hopefully we can continue that pace and get through the remainder of the measures by -- I believe we're scheduled to adjourn at 3:00, and we plan to do that.

So, we'll get started. We're going to start with the mental health measures.

If you look on your agenda, those are the measures on day two at 8:15 a.m. And we have -- we will be having some representatives from the American Medical

Association on conference call for the two measures, their two measures, 1364 and 1365.

We're a little surprised that your NCQA representative isn't here yet.

Oh, there you are. I didn't see you, sorry. Good, glad our NCQA representative is here.

So let's start with 1394, depression screening.

DR. ZIMA: All right. Good morning. This actually is a process measure at the provider level. It describes actually two things, that depression screening by 13 and 18.

However, I had some concerns about whether the description was consistent with the operational definition, because the denominator appears to be only for a visit within a one year window of the 13th or 18th birthday.

This seemed a little inconsistent with the text that said the target age range

was six to thirteen and thirteen to eighteen, so I think that will need a little clarification.

Also, the act of screening is not operationalized, and the rationale for the importance of this measure is really estimates -- prevalence estimates of depression.

I also wanted to, it was a little bit of a query, because some of the citations supporting the high impact of these measures include two research papers on ADHD and conduct disorder, prevalence estimate of serious emotional disturbance, which actually is based more on a functioning measure, secondary data analysis of mental health services among children and teens.

However, they do cite the USPSTF recommendation that screening only applies actually to teens ages 12 to 18. And this concept is supported by the AFFP, Bright Futures, and Michigan Quality Improvement Consortium.

The assumption, I think, within this measure is that improved detection will lead to improved care.

And I think, you know, over the last day we've had some discussion with other disorders of whether that's true and whether that fits the bar for a measure today.

I wondered whether improved detection really did lead to improved care, especially in privately insured behavior health carve-outs in Medicaid.

There was one study, however, cited, that showed that in a primary care practice about half of the teens that were detected got meds.

Reliability is not established, something that's been similar to some of the other NCQA measures proposed, and the validity is limited to face validity. Testing is not completed, so I wasn't sure if this was a time-limited measure, or --

DR. WINKLER: I think that we

ought to consider whether you feel the measure has been adequately tested for reliability or validity, and if not, then it would be a problem.

DR. ZIMA: Okay. In general, I think that there's more evidence supporting screening teens, but not the younger children.

And again, I think we need some discussion maybe to make a look at this one-year window versus how it's described with the age range.

And I think that was really about it on this one.

CO-CHAIR MCINERNY: I'm concerned about the value of screening under age 13, because I don't think there's really any evidence that that's going to be very helpful.

I certainly agree with the screening of the adolescents, since we've had the U.S. Preventative Services Task Force weigh in on that.

As far as treatment, I think

there's a little bit of hope there, because the Academy of Pediatrics recently released the -- what we call the mental health tool kit, and also published an article in, I think it was June Pediatrics, about the primary care pediatrician being able to treat mild to moderate mental health disorders such as ADHD, depression, anxiety, et cetera, often, and hopefully, in collaboration with mental health professionals.

And the tool kit is designed to give the primary care pediatricians some of the tools necessary to do that.

And the Academy of Pediatrics is really trying to get more primary care pediatricians into taking care of children with mild to moderate mental health disorders, since we know there clearly are not enough psychiatrists, psychologists, et cetera, to go around.

And I certainly agree with you about these mental health carve-outs, managed

care mental health carve-outs. They're poor for adults, and terrible for children. Hopefully we can keep those to a minimum.

So, I'm wondering, if we can sort of split this and decide on whether the 13 to 18 year olds would be -- decide on that, and then take a separate decision regarding the under 13 year olds.

DR. ZIMA: These are -- NCQA puts these forth as two separate measures, so you can act on them independently.

CO-CHAIR MCINERNY: Good. Thank you. Other comments?

DR. ZIMA: I'd like a second look on the denominator on that, and make sure that -- I was concerned about the inconsistency of the description of the measure, whether it was age range of 13 to -- even if we went with the other one through 18, or is it the one-year window of the child's 18th year? So, that's the first question.

And then I think the other

decision point is, again, sort of a question of the bar for the committee about detection, is that enough? Because there is no follow-up.

DR. GLAUBER: I have a question about the denominator. I know we weren't too kind to well child visits yesterday, but if the unitive analysis is the practice, and all that we're requiring for the measure is that there be a face-to-face encounter, and knowing that a lot of adolescents don't make a yearly well visit, then a practice may be held accountable for screening when the child has only been in for an illness or injury-related visit, and that's not something that's realistic or feasible from a practice perspective.

DR. ZIMA: And remember also, the suicide screening is not operationalized, so I wasn't sure, you know, if I was to get this data, what that would mean. Is it just a question, or is it --

DR. SCHOLLE: May I clarify the denominator choice?

DR. ZIMA: Thank you, yes.

DR. SCHOLLE: So, this is looking for children -- so the denominator are children who've reached the age -- the birthday, and had a visit -- during the measurement year, and had a visit in that 12-month period.

We're looking for screening to have happened within the measurement year or the year prior, so if it's children turning their 13th birthday, we're going back two years.

So they could have been screened during their age 11 and age 12. And we're looking at the Preventive Services Task Force recommendations, because we thought that that was consistent.

I thought the task force recommendations were age 12, so it would still work. But for the age 18 -- so it's by the

18th birthday. So screening from age 16 and age 17, it had to happen before their 18th birthday.

And I did want to talk about the numerator, definition of what is screening. In our field test, we actually looked to see whether a standardized tool was used, and we want -- because we did want to require the use of a standardized tool.

However, what we found is that there were -- there was evidence in the chart that depression was screened for, but rarely were standardized tools used. And so -- and the rates would have been extremely low using a standardized tool.

So, for these measures, our measurement advisory panel, this and the other mental health measures, our measurement advisory panel said, let's focus on screening, because they could be using -- you know, the two-item screener for -- and that might not be documented as a standardized tool.

DR. WINKLER: Question to bring up: if the folks from the outcomes committee recall, we evaluated the pediatric symptom checklist as an outcome measure, and as a change measure. So how do these all fit together, I think is a question to consider.

DR. ZIMA: I think just to clarify, the pediatric symptom checklist which the author was Mike Murphy and Mike Jelinek, the bottom line is that there's a lot more evidence that it improves detection of need for mental health services, but it was a time-limited endorsement because the data was not there to support it as an outcome measure.

DR. SCHOLLE: The Pediatric Symptom Checklist, the measures that were recommended by the U.S. Preventative Services Task Force were the PHQA and other measures.

So, and the Task Force recommended several different measures, and that's why we didn't tie this screening measure to a specific tool.

DR. ZIMA: Yes, in all fairness, the Pediatric Symptom Checklist is not just for depression.

DR. JENKINS: I'm still a little confused about the screening method that you're advocating for through putting this measure forward.

There was another measure developer who faced the same challenge, where they wanted essentially to push people towards a high reliability, high validity tool, but not force people to choose specifically one.

So, they actually specified in the measure that it had to be a tool that met certain validity and reliability requirements, and then provided a list of possible tools that a practice could choose to use.

And I must say, I don't know if that can be done here, but I'd like that much better than to say that when we did our chart review, we couldn't determine that an adequate screening tool was used from the

documentation. And our rates would be low if we had made that requirement.

So therefore, the measure we're putting forward didn't include that.

DR. ZIMA: Yes. And the approach of giving the psychiatrist or primary care provider an option of tools that are standardized is consistent with ADHD care.

CO-CHAIR MCINERNY: Kathy, and I'd like to second that motion. We're moving toward that in all respects. In developmental screening, we're not accepting that the pediatrician asked a few questions about the child's development and assumed their development was normal.

We want them to be using a standardized test, and I think the same should be true for the depression screening as well.

DR.CHEN: I'm sorry, so if that's the case, maybe we should review this one after? Maybe come back with a list of standardized tools, so at least we'll know

that we're comfortable with it or not.

DR. SCHOLLE: We actually have in our original field test specs, the standardized tool, and we have the language. So would you like for us to come back with a revised one that says, using a standardized tool?

DR. WINKLER: So, Sarah, when you were field-testing it, was part of the field test the use of the standardized tool, whichever one you've specified as -- and in the chart --

DR. SCHOLLE: We looked to see, was depression screening done at all? And then we looked to see, if yes, was a standardized tool used, and if yes, which standardized tool?

And we were looking specifically for the tools recommended by the U.S. Preventative Services Task Force, which are the PHQA and the BECK.

And I agree with you. It's very

rare that those standardized tools are used, and that is a better way to do case finding.

DR. ZIMA: Just a question. In your field testing, do you recall kind of the distribution of the age range you had, because some of the standardized tools for the older child might be more appropriate than the younger. Or else you have a very interesting data set that we should --

DR. SCHOLLE: So, we did -- so our field test was focused on those children turning age 13 and children turning age 18.

I think Sepheen's going to try to pull up the results very quickly, but my recollection of the data is that the use of standardized tool was rare, even in these high-performing practices that were part of our field test and also on the health plan side. So it's just not used very much at all.

DR. WINKLER: So, essentially, Bonnie, the opportunity that this measure could bring is to push the use of standardized

tools, and that would be a beneficial thing.

DR. ZIMA: Yes, I think that, you know, I think Sepheen actually taught me yesterday to emphasize baby steps.

And, as this little psychiatrist, I'll remain humble in my field today, how much we need to do.

DR. JENKINS: I think an alternative would be if you can justify why you think a high-performance metric should be based on a non-standardized tool. I guess we could consider that rationale.

But right now, it's hard to go all the way with that proposal.

DR. SCHOLLE: Our Measurement Advisory Panel felt that it was important to have discussion of this topic, and they wanted to recognize providers that were having a discussion of depression at all and documenting that in the children's chart.

So they wanted to -- they thought that was the first step, get people to start

asking questions and documenting discussions of depression, and they felt like that was the most feasible first step.

And the rates that we're looking at in these high-performing practices, 17 percent of the 13-year-olds there was documentation of standardized tool, 10 percent of the 18-year-olds.

But in the health plan population, we did a field test with health plans and with physician practices -- in the health plan, it was zero percent using a standardized tool. So we feel like that probably represents the real world.

So our panel, our Measurement Advisory Panel said, agreed with you on the developmental screening, standardized tool. And on depression, they said, you know, on mental health, there's really not a lot of consensus, and they felt like any kind of discussion and screening was the first step. That was the discussion.

DR. LIEBERTHAL: Did the participants express any thoughts about why they were not using standardized tools, and would they be willing to use standardized tools?

MS. BYRON: So in our de-brief, with the participants, they cited the fact that this task force came out, it's brand new. And they felt like there wasn't a lot of consensus around what the tools were.

Even across states, actually, if you look at Medicaid, they use -- a lot of them actually have -- they provide lists of tools to their plans or to their physicians, and they're all different.

So some of them are actually tied to what their state is doing, and they can't necessarily -- you know, if the state's going to pay for a certain tool, that's what they're going to use.

And so they're getting a lot of different information coming in at them, and

if a state is putting a tool on there that isn't necessarily on the list, and yet our measure has a different tool, they get caught in the middle. So that's why we wanted to just keep it to screening, documented.

DR. ZIMA: So, this is a very good conversation that I think -- and we don't want to be creating new measures in this process.

But I think what you're raising, which I'd be very interested in, is, as you look at the data, A, what percentage is there any documentation of even inquiry? You know, sort of a two-pronged provisional operational definition of screening. Was it asked? Yes, no? And then, B, was the tool used?

I think some of that operationalization there might be helpful to me to better interpret how this measure would be used.

MS. BROWN: A couple of questions. When you say, was it asked, I'm wondering if everybody knows what it is or what the exact

question would be.

I mean, I think there's a lot of variation in how people approach these issues, and I'm inclined to feel better about this measure if there were some discussion of these standardized tools.

So the question is, I don't understand. Are you all saying that there are screening tools that are well-studied, well-recognized, and can be used?

Okay, how many? Are there 12 of them, two of them, 50?

So, four or five. And are they suited to adolescents 13 and -- okay, so, can the measure, again, we're not talking about what happens if they test positive or we're concerned. I understand it's a screening tool.

But what's wrong with specifying either what the leading question is and/or what the range of useful tools might be? That really nails it down.

MS. BYRON: That actually is how we specified the measure when we took it to field test.

We did provide a list, and this is where we did get the feedback saying, your list is actually different from the list that we get from, you know, the Medicaid State Office, or your list is different from the list that we see here.

So I think there are a lot of tools out there. But when you start actually going through each tool, there isn't complete and total consensus around all of them.

MS. BROWN: Well, I understand. So what that means is that's a strong research recommendation, then, from this group, is to take the six leading state-mandated Medicaid ones and the four that the mental health -- and try and come up with the two or three that have the best evidence of effectiveness.

DR. PERSAUD: I suspect also that what's going on is that there are tools out

there that are covering a number of items beyond depression, and that's where you're getting into issues.

There are tools that have multiple constructs of things that they're screening for, or questions embedded in something else.

Like, we use a teen screen, and I think we have some depression questions embedded in there. You would not recognize right off that there was a depression screen in there.

And that's because of the issue of, for teenagers, trying to screen for all of these things as a practical issue.

DR. RAO: I think we're making the assumption that a tool is absolutely necessary to pick up depression, and that the tool has to be documented.

Now, I'm curious to see among your test practices, what did they find the prevalence of depression to be? Because if it's comparable to what we know nationally,

however they're screening for depression, they're successful. Do we have that information?

DR. SCHOLLE: We don't have data on the number who need it, although I believe it was a relatively small proportion.

DR. ZIMA: I'm having some empathy for the next meeting you guys are going to have.

And it's a problem, because the prevalence estimates of childhood and adolescent depressions are often based on the DISC, which is diagnostic criteria from the DSM-IV.

But when you're looking at some of these larger-based samples, they're more sort of generic screeners, like what Donna's talking about, which were really developed more to indicate need for mental health services, not a disorder.

So, and just one other side, I put on the list, particularly if we're going to

stick with some of these young kids, was there any discussion about the children's depression inventory?

DR. SCHOLLE: I believe we had that on a list. We really just didn't see any standardized tools being used. We did not bring to this committee a measure that we tried to do for age six for mental health symptoms where we had even greater problems.

DR. ZIMA: They actually are probably quite wise, given that the evidence for the diagnostic criteria is stronger for our teens than our younger children, and so I would imagine that was also a big issue in your measurement group.

CO-CHAIR MCINERNY: I'd worry about just relying on, the physicians said, I screened, because I think the screening could be, you're not depressed now, are you? Good, okay, let's move on.

And then they write down, yes, I screened for depression.

Well, clearly, that's not the recommended way to screen for depression. And so I think we need to raise the bar and say we need to use one of the several standardized tests to make sure that it was done appropriately.

DR. GLAUBER: And just speaking from a state in which, you know, screening is mandated by court order and it's paid for, it's pretty remarkable the level of performance that can be achieved just from billing data for 96110 for kids, we're up to two-thirds of kids that are being screened by a validated instrument.

DR. ZIMA: Do you have any data, what's happening with the kids that are detected?

DR. GLAUBER: Well, that's interesting, because we did our own internal analysis to see which -- what percentage of those kids, just based on billing data, wind up having a behavior health claim within 60

days of screening, because you can use a modifier to indicate whether a behavior health need was identified. And less than ten percent of kids had a behavioral health visit. But that's just based on claim.

DR. JENKINS: I just want to make one last point, as sympathetic as Bonnie is, to me, there's a bright line between a question you might find of value to ask, like, your very generic data that you got, and what you're putting forward as an official, high-stakes performance metric.

And I think that you're at the rate right now with some of your core questions where it's really just a quality metric. You're saying people should go and collect this data and try to understand it and understand what it means to take a baby step.

But to me, there is a bright line above that that is where this conversation is at, just to clarify.

CO-CHAIR MCINERNY: So I think

everybody's in agreement that we will defer this and NCQA will come back with some indication of what they feel about the standardized testing instruments.

DR. SCHOLLE: Okay, thank you.

CO-CHAIR MCINERNY: Good. Thank you.

DR. WINKLER: Do we have anyone from AMAPCPI on the line?

MS. TIERNEY: Yes, this is Sam at the AMA PCPI.

DR. WINKLER: Great. Okay, then we're going to launch into the discussion of measure 1364, child and adolescent major depressive disorder, diagnostic evaluation.

Dr. Persaud?

DR. PERSAUD: So, this is a measure of looking to assess how well those who diagnose children with major depression use the DSM-IV criteria. And it's looking for the percentage of times when major depression is diagnosed in children 16 through 17 years

of age, that at least five elements with duration of two weeks or longer, including either depressed mood, loss of interest in pleasure, how often those are used when the diagnosis is made.

This measure doesn't have, as put forward in the form, any contradictory information. It's recommended by evidence-based guidelines.

There is a significant performance gap in the data with most psychiatrists saying they either do not use the DSM-IV criteria or use it partially.

There's no disparity-sensitive information that I think this a measure that has potential for picking up disparities in care, and this proposes a population health measure.

But I think this is going to be, again, a quality metric and probably at the practitioner level of performance accountability credentialing metric

potentially.

It has been field tested. The field testing is face validity, and then some inter-rater reliability across checking charts, and just two things, I think, that are up for discussion, I hope Bonnie will have some interesting ideas on.

One is, there was an algorithm used for the numerator, which is extracting how many symptoms were used from the DSM criteria, and I'm not sure I could follow all of it, so it would be good to hear the measure developer describe that.

In one place in the form, I think they said it's administrative data, but I think that's going to be on the level of chart abstraction. And a flow sheet is mentioned, so I think there's some issues of feasibility of getting that numerator data. That's one.

And the other is that ages six through seventeen are all lumped together in this measure, and I do want to hear thoughts

about how closely the younger children fall into the DSM-IV criteria.

I think there are just two ways of looking at, one being stringent with that diagnosis, which is important, given the black box warning, et cetera, but the issue of sometimes having to stretch a little bit for the younger children because they don't completely meet criteria.

And those of us who voted were favorable to this measure.

DR. WINKLER: Just to clarify, Donna, this measure, if you look at this section, level of measurement of page seven, is for individual clinicians or groups. It's not population. It is at the practice level.

DR. PERSAUD: Okay, I missed that.

CO-CHAIR MCINERNY: Bonnie?

DR. ZIMA: I think again, the most important question is whether adherence to this measure is related to any improved outcomes. And the data's not there. I mean,

it may not be a fatal flaw, but I think it's still an issue that we wrestle with here.

The other thing, too, is just I actually am a little bit more skeptical about the strength of the evidence. I think the American Academy of Child and Adolescent Clinical Practice Guidelines have based more on literature review. It's not based on grading the quality of the scientific evidence like I think many of the people here are more familiar with.

And also, remember, the DSM-IV diagnosis of depression was initially created for adults, and then it sort of was adapted later for younger children with an irritability proviso.

But again, the scientific evidence for psychiatric disorders based on the DSM criteria is quite variable, depending on disorder, okay? So I think that tucked within this measure is an evidence base that is variable, okay, and that kind of is a little

bit of an Achilles' heel.

Also, I agree with Donna that the feasibility of abstracting this numerator is difficult, and most people do not document all five diagnoses for two weeks, disorder and persistence of at least that one target symptom of depressed mood anhedonia for two weeks.

So that also raises to me a little concern about the feasibility.

I also didn't understand the assumption that the implementation results are expected to be applicable to this pediatric measure, and a question I had was, was there any information on how the adult version performed in the CMS PQRI program.

DR. WINKLER: Someone from PCPI want to address any of these issues?

MS. TIERNEY: Hi, this is Sam Tierney again. Thank you for your thoughtful comment.

There were a number of questions,

so hopefully I can address some of them.

And I believe also we have one of the physician leaders on our work group on the line. Her name is Dr. Karen Pierce.

So, Dr. Pierce, if you have anything to add, please --

DR. PIERCE: I do. Can you guys hear me at all?

MS. TIERNEY: Yes, we can.

DR. PIERCE: Perfect. Here is my concern about the measure. You're right, it's a quality measure, because I think with the black box warning on the SSRIs and depression being a cause of significant morbidity in children, to put kids willy-nilly on meds without a clear diagnosis that meets criteria for depression, albeit adult criteria, it becomes a patient safety issue, as much as anything else.

So this is one of these quality measures where a clear diagnosis following guidelines is recommended because of the

patient safety data.

DR. WINKLER: It would be helpful if you could speak up just a little bit more. We're catching little bits of what you're saying.

DR. PIERCE: Me? I know I have a terrible line. I'm thinking about calling in. Can you not hear me now, or is it better?

DR. WINKLER: It's about the same, actually.

DR. PIERCE: Could I just call in again? Why don't I do that?

DR. WINKLER: Okay.

DR. PIERCE: All right. Bye.

MS. TIERNEY: In the meantime, this is Sam Tierney. I -- just to address a couple of the other questions that I heard, I know there was some question as to the scientific evidence and the quality, I guess, of the American Academy of Child and Adolescent Psychiatry Guidelines.

And I do have a copy of the

Guideline with me, and I'll just share with you how they say that it was developed.

They say that the treatment recommendations are based both on empirical evidence and clinical consensus and are graded according to the strength of the empirical and clinical support.

And they also offer a number of recommendations that are based and that are rated -- one of the recommendations that I think supports this measure is classified as a minimal standard.

And just for your information -- and I think this might have been included in the form that we submitted as well, according to their criteria, minimal standards are applied to recommendations that are based on rigorous empirical evidence and/or overwhelming clinical consensus. Minimal standards apply more than 95 percent of the time.

So I don't know if that addresses

those concerns.

Another question that I heard that was related to the feasibility of the measure, and I believe we submitted for you classifications for electronic health record extraction of the measure.

And many of the symptoms are available in codifiable field using SNOMED codings, so I think that from a feasibility standpoint we would be able to capture the data fairly easily.

CO-CHAIR MCINERNY: Thank you very much. To me, this is analogous to what we've done with the ADHD guidelines a while back in that we did tighten and say that pediatricians, if they're going to diagnose ADHD, they should follow the appropriate DSM-IV criteria for that.

And I think that's an important first step so that again, you're not putting children on stimulant medication who don't meet the criteria for ADHD.

And so I think this is a very important step for children and adolescents with depression.

I agree with you, Bonnie, that the holy grail of trying to get them all into treatment is not here, but at least we take this step first.

And again, the idea of making sure that only those who meet the criteria for depression are being put on SSRIs or referred for counseling I think is a very good first step.

DR. LIEBERTHAL: Bonnie, do you know if the DSM-5 will be more child-specific?

DR. ZIMA: No. And I actually went on the DSM-5 website to take a look and see whether there was going to be a big change in the diagnostic criteria, and I couldn't find that change.

But that is a question, because the DSM-5 will be scheduled to be published in 2013.

So a question I had for the NQF staff was really, giving the timing of your process, what are the implications?

DR. WINKLER: If an endorsed measure has embedded in something like DSM which will have an update, that would -- especially any changes to the DSM criteria that would impact the measure, that would prompt an ad hoc review that we could do at any time, because there's been change either in the evidence or in the guidelines that exist.

And that's true for all of our measures, so that doesn't have to be an overriding consideration at this point.

DR.CHEN: Can I just ask a question? I'm not familiar with this literature myself, but does the evidence support as young as children six years of age?

I mean, I would be comfortable with this if it is sort of mid to older teens, but I'm just not sure the DSM-IV criteria, the

categories would be fit for younger kids.

I think a lot of kids can present with depression in different ways, behaviorally or through anxiety or other disorders.

DR. PIERCE: Well, there is evidence -- this is Karen Pierce again. There is evidence --

DR. WINKLER: Can you speak up just a little bit? You're very faint.

DR. PIERCE: The concern is there is evidence that DSM depression criteria is valid in kids even as young as three and four, some of the preschool kids. So, it is a valid diagnosis if you are using the irritability criteria and using specific diagnostic criteria.

So it is important even in the younger kids, more importantly, to use the diagnostic criteria so it's not over-diagnosed.

CO-CHAIR MCINERNY: All right. If

there are no more questions, I guess we can take a vote?

DR. WINKLER: All right. So do we have any of the committee members on the phone?

DR. SCHWALENSTOCKER: Reva, this is Ellen.

DR. WINKLER: Ellen, I think I heard you, but you are so faint.

DR. SCHWALENSTOCKER: Yes, I'm here.

DR. WINKLER: Okay, great. Thanks.

Okay, so for the committee, how many feel that this measure 1364 meets the importance criteria?

Ellen?

DR. SCHWALENSTOCKER: Yes.

DR. WINKLER: Okay. So that's 14. Were there any no's? No.

In terms of scientific acceptability, how many feel that it meets the

criteria completely?

That's none.

How many partially?

Twelve.

Minimally?

There's one.

Ellen?

DR. SCHWALENSTOCKER: Partially.

DR. WINKLER: Okay. In terms of usability, how many feel it meets the criteria completely?

Two.

How many partially?

Ten.

Minimally?

One.

Ellen?

DR. SCHWALENSTOCKER: Partial.

DR. WINKLER: Thank you.

And feasibility, completely?

Okay.

Partially?

Minimally?

And Ellen?

DR. SCHWALENSTOCKER: Partial.

DR. WINKLER: Okay. Thank you. All right. In terms of recommendation for endorsement, how many want to -- yes to recommend the measure?

How many do not want to recommend the measure?

Are there any abstentions?

Ellen?

DR. SCHWALENSTOCKER: I would recommend endorsement.

DR. WINKLER: Okay. All right. So it's 11 yes, 2 no, one abstention.

Okay. So I guess the next measure to discuss is also from the AMA PCPI, 1365, child and adolescent major depressive disorders, suicide risk assessment.

Dr. Zima?

DR. ZIMA: All right. This, too, is a process measure at the provider level.

It assesses the percentage visits of patients ages six to seventeen who have a clinical diagnosis of major depression and have had an assessment of suicide risk.

The main rationale for the importance of the measures is that they argue that major depression is prevalent and that suicide is the third-leading cause of death among teens.

The citations supporting the opportunity for improvement, however, appear to be based on adult samples, and I think the strongest argument for this measure is that suicide is a bad thing.

Its scientific acceptability is that it does not provide evidence for assessing how actual -- how assessing the risk reduces the risk of suicide. However, we're always taught clinically that you're supposed to do that.

However, like some other measures I think we've discussed, this measure, too,

specifies something that you're expected to do clinically, and frankly, you would fail your boards in psychiatry if you didn't.

The method for rating the strength of the recommendation did not rate the quality of the scientific evidence. And again, just like the earlier -- indicated that we talked about it, it was really based more on clinical consensus where the academy -- that it's something that you should do.

I think the assessment for suicide risk is not operationalized in this measure. And again, I'm not sure whether that's something that people feel strongly that there should be in there, or how you would interpret screening for suicide risk.

The denominator only includes youth that have a clinical depression, so I think it's important that in this measure, screening for suicide risk would be dependent on first detection of major depression in the child.

As far as reliability, it appears that the testing that's noted is not really directly related to this process measure. Validity is based only on face validity. There's no information provided for risk adjustment.

And again, this might be again, I understand, premature for the field, but just to keep in mind that the clinical risk of suicide is much higher among older youth and also other risk factors for suicide that you would take into account confound with age.

So things like substance abuse and access to weapons, driving a car, would all be confounded with age. So I think it needs to be said.

As far as unintended consequences, just one concern was if we detect suicidal ideation but this measure does not require any follow-up, what do we do with that?

What are the legal implications for a physician that documents -- and that's a

very thorny issue, I understand. But what would be the reporting duties of the physician if he or she documented that in the record?

And then I think, again, the rationale for the implementation of this measure is based on the adult version, and at least here we had some data that among adults, the percentage was 81 percent in 2008.

DR. WINKLER: Any other comments from any other committee members?

Does anyone from PCPI want to make any comments in response to Dr. Zima?

DR. PIERCE: Other than her comments, this is Karen Pierce again.

DR. WINKLER: Okay. All right.

DR. LIEBERTHAL: It was mentioned that the assessment of suicide risk wasn't operationalized. I assume that means that there were no guidance as far as structured evaluation. So what constitutes an assessment for suicide risk in a general provider's office?

DR. PIERCE: I think even asking, it turns out, if you look -- this is Karen Pierce -- if you look at data from ER samples, if a kid shows up in the ER, is suicidal, parents, doctors, and kids all know that they're suicidal, two weeks later, you evaluate the kid.

The kid is still actively suicidal, but often the doctor and the parent do not know because the kid is looking better.

So this is a safety measure, you're right. There is not much evidence, but the evidence that when kids start looking quote "better," they're often at higher risk.

And it needs to be just alert and above, because most suicides, a patient has seen a mental health provider within the last three weeks and has not been queried about their safety.

DR. JENKINS: I have a clarifying question, then. Is this measure intended to be used by the psychiatrist or by the

pediatricians or by any health care encounter that I guess that a child with major depression has?

I'm a little confused about it. And I did think there were some validated instruments, for example, with ER encounters and inpatient encounters about screening tools for assessment of suicidal ideation. I don't know if that's what you're referring here to or to something different than that.

MS. TIERNEY: This is Sam Tierney. I would say that this measure is broadly applicable to any physician caring for patients, child and adolescents with major depression. It would apply to primary care physicians as well as psychiatrists, and others treating these patients.

MS. BROWN: This is Sarah Brown. I'm still confused, though, about what the question is, or what the tool is. Is this any question, or a set of questions? What actually is -- what is somebody supposed to

do? What is a provider supposed to do?

DR. PIERCE: You're correct. They're supposed to just ask. There is not a tool.

DR. WINKLER: We can't hear you. Can you speak up?

DR. PIERCE: You're right. There is not a tool. There is one suicide tool, David Shaffer's tool, which has been validated, and is a nice questionnaire. And his data suggests that's good.

But this measure does not recommend a tool at this point. It's right now a clinical discussion.

DR. JENKINS: Just to make a point, we just had another measure which was in the same situation. And I certainly saw it at this level. It sounds more like a quality metric than a performance measure, because it's not well specified or well validated.

MS. TIERNEY: This is Sam. If I could just add -- I mean, I think the work

group plus the type of assessment brought on purpose, you know, this is required at every visit.

And so the specific type and magnitude of the assessment required by the measure is intended to be at the discretion of the individual clinician and should be specific to the needs of the patient.

So maybe at an initial visit, you might do a different type of assessment than you would when a patient's receiving ongoing treatment.

Or as Dr. Pierce mentioned, if a patient's gone to the ER, you might want to do a different assessment subsequent to that visit to the ER.

So I think on purpose the measure is left broad, because it really should vary and be specific to the needs of the patient at that time.

DR. WINKLER: This is Reva. Question. You have indicated that these are -

- you have EHR specs for this particular measure. Is there a particular location this information is going to reside in an EHR? Because with it sort of vague like that, it seems like it might be not necessarily found in the same place in an EHR. Have you further specified where the data will reside?

MS. CHRISTIANSEN: Hi. This is Kerry Christiansen from the AMA, and I will be happy to address that.

Along with many of our measures, the way you document information in the electronic health record does definitely affect the feasibility of it. But we feel it is possible to design the way you do your documentation around this to be able to collect the data in a standardized way.

Does that kind of answer the question? The feasibility is dependent on the design of your electronic health record.

DR. JENKINS: Is this a time-limited question, or is this a full

endorsement question?

DR. WINKLER: I think they said that this has really not been tested, and I don't believe these EHR specs have been tested, so I would say it's a time-limited.

Okay. Is everybody ready to -- Bonnie, did you have something more?

DR. ZIMA: You know, this is a difficult one for me, because in psychiatry, it's a medico-legal issue. You absolutely have to screen for suicide in any child that's depressed. I mean, it's -- so.

CO-CHAIR MCINERNY: Okay, shall we vote?

DR. WINKLER: Okay. How many on the committee feel that measure 1365 meets the importance criteria? Yes?

Any no's?

Ellen?

DR. SCHWALENSTOCKER: Yes.

DR. WINKLER: Thank you.

All right, in terms of the

scientific acceptability of the measure, how many feel it completely meets the criteria?

That's zero.

Partially meets the criteria? One, two, three, four, five.

Minimally meets the criteria? One, two, three, four, five.

Doesn't meet the criteria at all? One, two, three.

Okay. And Ellen?

DR. SCHWALENSTOCKER: Minimally.

DR. WINKLER: Okay, great, thank you.

In terms of usability, how many feel it meets it completely?

That's zero.

Partially? One, two, three, four.

Minimally? One, two, three, four, five, six, seven.

Not at all? Two.

Ellen?

DR. SCHWALENSTOCKER: Partially.

DR. WINKLER: Okay. And now feasibility. Completely meets criteria?

That's zero.

Partially meets criteria? That's zero.

Minimally meets criteria? One, two, three, four, five, six, seven, eight, nine, ten, eleven, twelve.

And not at all? I see one.

Ellen? I can't hear you.

DR. SCHWALENSTOCKER: I'm sorry. Minimally.

DR. WINKLER: Thank you. All right.

So, recommendation for endorsement, and this would be a time-limited endorsement, who says yes? One, two, three, four, five, six.

How many no's? One, two, three, four, five, six.

Any abstentions? One.

Ellen?

DR. SCHWALENSTOCKER: I think yes for time-limited.

DR. WINKLER: Okay. Very close, with seven yeses, six no's. So I think there's some issues on this that -- are there any follow-up you would like, additional information from the developers? Anything? Because an almost split vote on the committee isn't an enthusiastic recommendation. So is there something that would help?

DR. LIEBERTHAL: More of a specification of what screening for suicide risk means.

DR. JENKINS: And more of a specification in certain settings and certain timeframes by certain individuals, as well as the issues about where -- about the documentation and how the audits will be done to make the assessment.

DR. ZIMA: Yes. It was a little premature, I think, given the evidence base and the lack of testing.

DR. PERSAUD: I think if there's any other language or evidence that can be found out there to support its usefulness, that would help just a little bit more.

DR. WINKLER: So, folks from the PCPI, you've heard the concerns of the committee, and we will come back to you and see if you can address some of those, and then bring any additional information back to the committee for them to feel -- to look at.

CO-CHAIR MCINERNY: Very good. Thank you very much.

DR. WINKLER: I guess the next one in this group is measure 1406, risky behavior screening. This is a measure from NCQA and Sarah Brown.

MS. BROWN: Well, I think this is familiar territory to us now, which is a very important set of issues. I mean, a lot of what ails adolescents or burdens adolescents is in the behavioral domain, so I don't think there's any question that the issues described

here are very important.

I think our challenge is, what do we feel about the evidentiary base or the value of asking about these questions and counseling a referral? I think they probably vary enormously between the different issues described.

Just for example, there is a fairly well-developed although imperfect system for handling family planning and contraceptive services for sexually active adolescents. I'm not sure that applies to all the other issues that are described in this.

So I really -- without going into all the details, I think the question is, what's the yield here? It may be that we say, you know, this is such as important domain of issues that simply asking about it is again, one of those baby steps.

But I don't think we can say that there's a lot of evidence that particularly in busy practice settings, busy clinics, that

sort of ripping through a list of asking young people about this stuff is going to make an enormous difference. And the materials freely admit that this is based on expert opinion and consensus and there are very few data cited.

I also, just because it's going to come up again and it already has, I'm still confused. I'd like the NCQA people to just explain to me what the 19 physician group survey was. That seems to be a basis for a number of these.

And I don't know if that's a large and robust group, highly varied, lots of diversity in patients and practitioners, or really what that is and how sort of strong a base of data that is for any of these measures. There are several others that have used this same system as well.

So, without going on, that's sort of the overview here.

DR. SCHOLLE: May I respond, then? First, I'd like to clarify that there are

four different rates here, just to make sure everybody understood that, that we're looking at screening for each of these four activities.

The physicians, we worked with the American Academy of Pediatrics Quality Improvement Network to identify physicians who are willing to participate in this field test, and the field test was designed to test the composite rather than individual indicators.

So each of the physicians had about 20 adolescents that were represented out of the 50 children that they conducted chart reviews.

So we gave them detailed specifications and asked them to provide us data on each of these measures. So we said, tell us how many -- identify a sample -- or a consecutive sample of children aged 13, starting from a particular date.

And then they told us, was there documentation of a discussion of these risky

behavior topics.

Now, at the same time, we did field testing with five health plans. And actually, it's really -- those two sets of data that our panel was reviewing, because the physicians, we know are people that are part of a quality improvement network, probably more attuned to these kinds of topics.

The health plans we think probably represent the usual practice better in this. And the health plans, those were also chart reviews conducted by health plan staff.

The performance rates actually on this measure -- as in all the measures, the performance rates were higher among the physicians than among the health plan -- than the children that were identified through the health plan sampling.

But we tended to see, you know, that there was more discussion of these topics with -- or documented for the age 18 and the age 13, but probably about a 50 percent

response rate for the age 13 on each of these topics.

MS. BROWN: Sarah, can you also comment on the actual structure of the questions asked?

I mean, how were these -- there are four different ones, and they all -- again, standardized tools, or just, did you ask about x?

DR. SCHOLLE: For this, we looked to see whether there was any note indicating the date, and whether the provider asked or counseled about the following topics. So that's what the directions were. Did somebody ask about it, or did they counsel about these topics, sexual activities? So, that was as much direction as we gave.

MS. BYRON: Well, there's a little more. I mean, we also noted, you know, we gave some guidance as to what would count as the numerator hit.

So we said, documentation of

counseling could include a note indicating one of the following: engagement in discussion of current risky behaviors, checklist indicating that risky behavior was addressed, counseling or referral for risky behavior education, member received educational materials on risky behavior, and anticipatory guidance for risky behavior.

Now, these are based on existing measures where we do also ask for counseling. So we found that structuring it that way sort of gives them enough guidance, and our auditors enough guidance, to see whether or not a certain documentation would qualify as the person having received counseling. So those are modeled on specs that we have found to work in the field.

DR. SCHOLLE: And the other thing is, we looked for both a checklist where somebody's reporting something or counseling in the absence of documenting the behavior, because one of the things that our measurement

advisory panel talked about was a privacy concern for the teens, and whether you'd have this discussion, but you wouldn't document the results.

We actually found that the results were often documented, but we wanted to give practices credit where they were having a private conversation with teens, not documenting the results in the chart to protect the teens' privacy.

DR. PERSAUD: I have a couple comments. I think that I'm less concerned with this measure about how they actually looked at the chart to find out whether they were asking questions, because I think that there's many people practicing medicine where they look after adolescents have long since been overrun by what it takes to assess them, and many of us use questionnaires that cover these questions.

And this is why we can't pull out the depression, because what you're going to

find is most of them are 20-to-40-item questionnaires that they're using. So I'm less concerned about that.

I think the issue about the usefulness of whether, when we pick up things in the adolescents, what's our ability to affect outcome. This might be akin to the discussion we had about obesity and the BMI.

So, I think that we do have to screen the adolescents. We can't get out of that. It's just that, what is the ability of the primary care practice to affect an outcome?

And I think there's a good reason why it's better in the area of STIs and birth control, because you can prescribe birth control. So we're definitely going to do that.

What the issues is, is when you get to the substance abuse issues, those are only going to be amenable to moderate and intensive counseling, except for very mild

circumstances. And those would require referrals.

So I think that's where the evidence breaks down, because the addiction medicine field right now I think is not performing.

And I guess I speak in part of a parent of a child with mental health illness and addiction, what it took to deal with that. I think the problem isn't screening. I think we shouldn't measure screening. I think the development of what is effective for substance abuse hasn't occurred yet.

CO-CHAIR MCINERNY: For better or for worse, about three or four years ago, the New York State Department of Health and Medicaid, et cetera, did adopt this very measure, and is looking at charts, reviewing charts to see if the physicians did screen for these behaviors.

And they felt it was important enough to do, given the fact that the

evidence, I agree, is somewhat sketchy at best. But at least that's what's happening in New York State.

MS. BROWN: A few minutes ago, though, you made a comment about just screening for depression, just saying, you're feeling better, or how are you doing today, and then it's documented.

I think that applies in equal measure to this set. Without any guidance -- I mean, unless people are using checklists, and I understand that, but probably a number are just saying, are you sexually active? -- yes, no. Do you -- you know, my daughters get asked all that. Do you use illegal substances? And then -- you know.

So, I, you know, I think this issue of actually what it is as for suicide risk or depression matters a lot.

On the other hand, again, I think this area is so important that there may be value just in asking, if nothing else, to

educate the provider that these are really critical issues.

And hopefully, when there's more data on prevalence, it increases the pressure to develop better community systems to respond and provide care and treatment. That's the best case argument, I think, to be made for this measure.

DR. SCHOLLE: We did try to find whether there were any standardized tools for these, and what we understood was that there's some sort of acronym that's often documented in the chart?

DR. PERSAUD: HEDS interview.

DR. SCHOLLE: Right.

DR. PERSAUD: And that's why the tools are not easy, because most of us are using a combination of the tool and then we have to deal with the report generation with the adolescent, so it is harder to define.

DR. WINKLER: Donna, that brings up just a technical point. If indeed people

are using checklists where you're asking questions that will address several of these measures, whether it's risky behaviors or depression or this or that, and those checklists and questionnaires might be different among different practices and practitioners, how easy is it to reliably abstract that data on chart review? Do the abstractives know where to look?

DR. SCHOLLE: We got a lot of complaints in our field tests, but this is not one of them. But I think what we did find is that the teens -- that in some practices, they had the questionnaires, and then it was pretty routine. You could go in and find the questionnaires, or it was documented in the visit notes.

So I think that's the challenge. Our goal is to move this into EHR specifications and to think about that.

But I do think both documentation that comes directly from the child and

documentation that the practice team might use would count here. That's what we wanted to do is to allow both of those to count.

DR. JENKINS: Was this being put forward fully tested, or is this a time-limited question?

DR. WINKLER: I mean, Sarah and I have talked about this, and this group of measures have been tested for feasibility, but they really have not been tested for reliability or validity beyond just assessing the face validity.

DR. JENKINS: So just to make the point, reliability and validity would be that it captures the patients that it should capture based on how the chart audits are classifying whether that patient was screened or not.

And that's a much, much higher bar, but that's actually what we're talking about here. I don't think this has been fully tested.

CO-CHAIR MCINERNY: Do any of the practices use the guidelines for Adolescent Preventive Services, the GAPS questionnaire?

DR. SCHOLLE: I don't know.

CO-CHAIR MCINERNY: Because that is one that's been around for quite a long time, and is felt to be reasonably reliable.

Of course all of these things require that the adolescent is truthful, and we know, guess what, that's not always the case.

We also, though, tend to know that they are a bit more truthful when they're filling out something on paper than when they're being asked the questions face to face.

DR. SCHOLLE: So in the specification, just to be clear, it allows either the checklist was used, or if there was any counseling provided or documented.

And we found that a number of the practices did use questionnaires, and so that

counted. But also, if there's discussion in the absence of the documentation of what the child's response is, then that counted as well.

DR. ZIMA: It's interesting. I don't think this issue obviously is going to go away, and one of the things that I hope that you'll be discussing in the measurement is, maybe in the next year or two, when we can have teens actually enter the sensitive information directly into computers in the waiting rooms, this will capture even more Dr. McInerny's issue of trying to get more truthful information from our teens.

DR. SCHOLLE: That's actually what we proposed to do in our pediatric quality measures measurement program application.

CO-CHAIR MCINERNY: Well, you know, certainly the Academy of Pediatrics recommends that starting at about age 12 or 13 that the physician in the room with the patient alone asks the sensitive questions and

not with the parent, and that's a good first step.

However, I don't think that the adolescents truly trust that that will keep sensitive information from reaching the parents. And one of the biggest problems is that if there's any laboratory testing or referral done, then when the insurance company pays the bill and it goes to the parents, there's an explanation of benefits.

And the parent is going to say, hey, wait a minute, how come you had a test for an STI, or how come you got referred for counseling for this or that? And that's a problem, and it's a very difficult problem to get around.

MS. BROWN: I just want to underscore that a million times over. If there's any notion that the parents are going to find out about sexual activity, drug use, it's a dealbreaker in many, many instances. And I don't think that's fully addressed in

this.

I mean, it's just a reality of life. And it's not your responsibility to solve that problem --

DR. SCHOLLE: And that's why we --

MS. BROWN: -- but it really limits the value of this.

DR. SCHOLLE: Well, that's why we allowed that documentation of discussion, because we really --

MS. BROWN: Right.

DR. SCHOLLE: -- wanted it to be brought up and if -- to allow that discussion to be documented without having the results documented, in case that's what was happening. It's a challenge.

CO-CHAIR MCINERNY: Ready to vote?

DR. WINKLER: Committee ready to -- so how many of you feel that the importance criteria have been met for measure 1406? One, two, three, four, five, six, seven, eight, nine, ten -- Ellen?

DR. SCHWALENSTOCKER: Yes.

DR. WINKLER: Thank you. And were there any no votes? Okay.

In terms of the scientific acceptability of the measured properties, how many feel it is met completely? Zero.

Partially? One, two, three, four, five, six, seven.

Minimally? One, two, three, four, five, six.

Not at all? No. Okay.

Ellen?

DR. SCHWALENSTOCKER: I think partially.

DR. WINKLER: Okay. Thank you. Usability, how many meet completely?

Partially? One, two, three, four, five.

Minimally? One, two, three, four, five, six, seven, eight.

Not at all?

Okay. Ellen?

DR. SCHWALENSTOCKER: Partially.

DR. WINKLER: Thank you. And feasibility, completely meets? Zero.

Partially meets? One, two, three, four, five, six, seven. Okay.

Minimally? One, two, three, four, five, six.

Ellen?

DR. SCHWALENSTOCKER: Partially.

DR. WINKLER: Okay. All right, so recommendation for endorsement, and does everyone agree with Kathy that the measure is really a time-limited requirement? Okay.

Recommendation yes, for time-limited, got it.

Time limit is endorsed only for two years -- oh, I'm sorry, you're right. It's just been changed to one year. But does it come with?

DR. JENKINS: But the burden is on the measure steward to fill in all of the measurement issues whether it relates to

reliability, validity, and testing.

DR. WINKLER: Right, exactly. Any no votes?

Ellen?

DR. SCHWALENSTOCKER: Yes.

MS. SCHOLLE: Just to be clear, the time-limited endorsement means that given a one-year time period, it would be very difficult for us to come back with a revised measure with more reliability and validity information within that time period, because we don't have money right now to go do it immediately, so just so that you're -- the Committee's aware.

DR. WINKLER: Yes. Well, and something that is part of the ongoing follow-up is, any of these measures that are recommended for time-limited, one of our jobs is to go back to the measure developer and talk to them about the requirements for that before that can go forward.

So the fact that that's where

you've established yourself, we can do that follow-up.

CO-CHAIR MCINERNY: Can you provide an extension for another year if it they say yes, we're working on it?

MS. BOSSLEY: The short answer is probably no, not another year. We're getting significant concerns raised by the Consensus Standards Approval Committee and the board with having time-limited measures out there without testing it for reliability and validity.

But what we can do is work with NCQA and see if there's some in-between, something we can do. Yes.

DR. WINKLER: We are expecting another measure to vote for on the phone at 9:45.

Do you think we could maybe squeeze in one measure from yesterday before then? Okay.

And that might be -- how about

1390? This is child and adolescents access to primary care practitioners, measure from NCQA, Dr. McInerny, that was your measure from work group three. Measure 1390.

CO-CHAIR MCINERNY: Sure, I'd like to -- let me try and grab that, get that up here so everybody else can get a chance -- because that's a different work group.

DR. WINKLER: While everybody's grabbing it, I've pulled it up. The description of the measure is the percentage of members 12 months to 19 years of age who had a visit with a PCP.

The organization reports four separate percentages for each product line. Children 12 to 24 months and then 25 months to six years who had a visit with a PCP during the measurement year, children 7 to 11 years, and adolescents 12 to 19 years who had a visit with a PCP during the year or the year prior to the measurement year.

CO-CHAIR MCINERNY: Yes. And this

is really an access measure. It's not looking at whether this was a comprehensive well child visit but rather just was there access to the primary care physician in those various time periods with the notion that it is important to have access, and if there's no access at all, then that's clearly a problem.

So I think for that reason, when I reviewed this I felt that this was an important measure.

And then if we look at the scientific acceptability, that was pretty straightforward, and I felt that that did meet the scientific acceptability completely.

And then when we -- if we look at the usability, again, to me this seemed to meet the usability very completely as well.

And then feasibility, this is a pretty easy measure to measure because it's using administrative data. And so I felt that this did meet completely the criteria for feasibility.

I don't know -- frankly, I did not see what the other members who looked at this, what they said.

Okay. The voting was, impact, 60 percent completely, 40 percent partially, gap, 80 percent completely, 20 percent partially, and relationship to outcome, 80 percent, 20 percent.

DR. GLAUBER: Question. What are we left to infer about the children who don't satisfy this criteria in terms of access?

And wouldn't this be something that's better assessed through survey data about having a usual and ongoing source of care rather than just one face-to-face visit?

DR. RAO: I just wanted to echo what Jim said. I think that there is a lot of episodic care that takes place in primary care settings, so it would be nice if we had two visits as a standard in the same calendar year. That would at least demonstrate that there's some continuity there.

DR. GLAUBER: Or tying this, as we discussed yesterday, to having an ER visit and in the absence of primary care visits.

DR. LIEBERTHAL: Where would it fit if the patient has a same-day urgent visit with a pediatrician, but the pediatrician does not have a primary care practice, such as in a -- we have urgent care within the department, so they could have a couple visits with the same doctor. Hopefully that doctor's going to update immunizations, but it's really not a primary care visit, so I think there's a lot of gray zones in this.

DR. GLAUBER: Or on the converse, the kids who don't satisfy this, I'm not sure that we can infer that they don't have access to care. It may be a reflection more of their health status or the measurement period being too short.

CO-CHAIR MCINERNY: Well, with the proliferation of urgent care centers and retail-based clinics, there is an erosion of

patients visiting their primary care physician for illness visits.

And I think it is important to document what's happening over time about access to care to the primary care physicians.

The American Academy of Pediatrics, frankly, takes a very dim view of patients going to retail-based clinics and urgent care centers, because they're not staffed, in most cases, by folks who really have a thorough knowledge of pediatrics in the first place.

And secondly, there is an argument, and I'm not sure how -- although, this is a nice argument, there's no evidence to back it up, that presumably, if a pediatrician sees a patient for an illness visit, they might notice that there's some other kind of a chronic or an ongoing condition that needs to be addressed, and addresses it at that visit, which would likely not happen at an urgent care center or retail-

based clinic.

DR. RAO: I just wanted to ask, maybe Reva would know this, but do we have measures already that talk about whether, did we approve them, that children have a medical home, or an ongoing primary care provider?

DR. WINKLER: We have, in the outcomes project, you recommended the survey-based measure of the medical home.

DR. RAO: Okay.

DR. WINKLER: And so that is one measure there. Otherwise, I don't believe there are any others.

MR. STEINHART: This is Amos Steinhart. I joined the call a little bit early.

DR. SCHOLLE: Just to clarify, emergency visits don't count towards this measure. It is ambulatory, and I'm not familiar with the billing codes that are used by emergency centers, and whether they would do that. But it's not the intent of this

measure to capture urgent or emergent care.

And the value or the benefit of this measure compared to a survey measure is that it captures the entire enrolled population, as opposed to surveys that are usually done on people that have, you know, a sample of patients. And generally, we have very low response rates, and one might expect that there might be some non-response bias.

MS. CARLSON: This is a health plan measure, so ideally, the measure is supposed to summarize the ability of the health plan to provide access to their members to necessary primary care services.

I don't know if it does that. I mean, it's truly just a descriptive statistic. And I'm not sure that you can make a lot of inference from it. But it's a measure that's been around for some time. So you can see trending with it. And health plans use it in different ways, and they do use it when they're looking at access and their services

and trying to drill down to specific areas where they may have difficulties.

DR. JENKINS: Carol, I agree. And I guess what I would state it as, I'm struggling with it, as I was with many of the similar measures yesterday as an accountability measure for the plan, as a quality measure, you know, for the plan to track and try to understand, I totally understand this measure. As a population health measure, I understand it.

But as an accountability measure for the plan, I don't think it's a good measure of access. People might have access that they choose not to take advantage of it or don't feel that they need.

DR. LIEBERTHAL: I think the issue is, in the broad term, PCP, and I think it might be better defined as within a health plan, the child should have an assigned primary care provider.

And I think then to have at least

one visit with the assigned primary care provider rather than just anybody who could qualify as a PCP, whether they have any relationship to the child or not, how to phrase that, I'm not sure.

But that's the direction -- that's where I see the value of this, and this relationship to access.

DR. SCHOLLE: Just to clarify, most health plans do not require patients to have an assigned PCP. They encourage it, and they encourage people to change their PCP if it's not working out for them, so that would actually be very, very difficult to operationalize.

DR. PERSAUD: I think also that even then, there, I think what this is boiling down to is that it is hard to decide who it is that is practicing the scope of primary care, and that's what's hard.

You'll have a bunch of middle levels or primary care pediatricians, but they

could be doing different scope of practice care.

I think this would maybe over-generalize and maybe make it look like there's more access when there could even be less.

There are children certainly that turn up in primary care practices where the scope of care delivered is primary care, but they come in and they don't engage in primary care.

They come in seeking urgent care and get that and are told, come back for a check-up, and they don't return again.

So, as Goutham said, we tend to, in the primary care services at practice level, we tend to look at two visits as a more ideal way of assessing whether they're really accessing primary care.

And I think, I mean, if that's the whole sentence, it's not just access to care. Are they accessing true primary care?

I noticed under the testing and

analysis that there wasn't any reliability, and the validity was expert opinion, so I wanted to know if the measure developer had any comment about that.

DR. SCHOLLE: So, these are claims-based measures, and this is not a measure where we can provide comparison of the chart data to the claims, because we really do see the claims as the place to find information about visits.

We tend to use a chart review data to go back and document -- validate information like diagnoses. And in terms of reliability, we can provide you plannable reliability information, but we don't have it with us today.

DR.CHEN: Sorry, can I just comment on a few things? So I actually agree with Sarah in the sense that I have a lot of experience with survey data, and also with claims data.

And I think if you're interested

in looking at access, there are some things that survey data does better. But I think in this case I'd rather have hard facts, that they actually went to a visit.

So, in that sense, I think this is the right data collection approach, rather than using survey data.

Now, the issue is, are we interested in this measure as sort of the standard, very crude access measure, where they can just satisfy so the health plans can monitor their access and then their use and utilizations of services?

Or are we more interested, in, like Allan's saying, that we actually wanted to raise the bar a little bit where we really want them to see a primary care provider, whether or not they provide primary care or not or well child care, I don't know if we can discern that from administrative data, but at least we wanted them to see a primary care provider.

And if we do endorse it that would then, it is up to the health plans to make sure that they're assigned to a primary provider.

And I have done a lot of study on the MEPS data, where you could actually do this in actually administrative data where you know if they are assigned to a primary care provider, and you know if they had a visit to that primary care provider for that survey year.

So I mean, I think it's possible. I don't know if it's easy. But it's certainly something that if we think it's important for a kid to see a primary care provider during that period, then that's how we should try to push the bar that way.

But as a purely access measure, I think it's really an absolute minimum in this case.

DR. GLAUBER: But one question I have, in terms of it as an access measure, do

we know what percentage of kids -- you know, I'm thinking of kids with special health care needs who may be exclusively cared for by a sub-specialist or a behavioral health clinician who have access to care but are not seeing a primary care doctor, so it's really a coordination of care issue rather than access to care.

DR.CHEN: Right. I mean, that's a very good point, although I think I would say most medical home agencies or proponents would promote that in the medical care team, there would be a primary care provider that's the general pediatrician that would be the team.

I mean, even though they don't have to have visits with that general -- and that happens to a lot of my patients, actually. They see cardiology as their primary care, but I would still feel comfortable that they come see me because I provide the care coordination, make sure everything goes well. So, I mean,

especially for those kids. I mean, they have multiple visits in a year. I don't think having a visit to a primary care provider is an extra burden to them, but I don't necessarily think that primary care should be provided only by general pediatricians, so I agree with that.

CO-CHAIR MCINERNY: Well, certainly one of the cornerstones of the medical home is access to care.

And when you look at some of the measures of medical homeness, they look at, how does the physician practice work to make access to care easy? Do they offer same-day appointments, after-hours appointments, et cetera and so forth?

And so in that regard, I think the access measure is an important measure, true at the health plan level. But I also think that it's an important measure at the medical home level as well.

DR. JENKINS: I'm still a little

troubled with the fact that yesterday we talked about well child visits for well children in the adolescent age range on an annual basis and found that problematic from an evidence perspective.

And this is a measure between age 12 months and 19 years, and it's asking essentially the same question, in the case of a well child, someone who did not think they needed to go to a doctor and chose not to go.

So I'm troubled by the disconnect there, and I just need some help with it, along with the accountability issue.

DR. SCHOLLE: It is different from the well child visit, because it does not look for the specific CPT codes that are for well child care. It's for a visit with the primary care physician, so it's access to the primary care physician, so it's not well child.

DR. JENKINS: Well, I guess just to make my point, and I'm stuck on the patient who chose not to come in because they did not

feel they needed medical care.

The plan -- that patient will be in the plan's denominator, and not in the numerator. So let's suppose 20 percent of the adolescent patients chose not to do that. They get a grade of 80. That's their grade. Okay?

And it's reflective of that 20 percent who chose not to come in. And I'm still struggling of, should they have come in?

I mean, are we saying they should have come in, the plan should have brought them in? How could the plan have brought them in to give them a grade of 80 instead of 100 on this measure in the adolescent age range? That's my problem.

DR. LIEBERTHAL: I think that the intent of the measure is very, very good. My concern is that the definition of it doesn't achieve the intent, and I'm just off the top trying to figure out how to make the numerator statement achieve the goal.

And I can't do it right here at the table, but I think that more effort needs to go into a better definition that would achieve the access goal.

CO-CHAIR MCINERNY: Well, this would change the measure significantly, but one way to get at that would be to see if the patient went to a retail-based clinic, urgent care center, or emergency room in that same period of time.

And if they went to any of those three or four times, and not to the primary care physician at all, then that seems to me that's an access problem. And I would pull out -- sure, I suspect there are some adolescents who feel they're pretty healthy and they never go anywhere for anything.

And you know, one can argue that, appropriately, that that may not be fair to ding the plan for that reason.

MS. CARLSON: Well, I think that gets back to my comment before where you

really need to take some of these measures and look at them together to try to determine what the actual outcome is.

So you really have -- this is one of those measures where, as a health plan, we would look at it along with several other measures in a market basket approach, and decide whether or not we have a true access issue.

We would take this data and compare it to our data for preventive visits, along with maybe data about how many physicians we have in network within certain geographic areas. And we'd look at that together to try to determine if we have an access issue.

So I'm concerned that maybe our expectations may be too high based on the use of one measure as opposed to using a number of measures to get at what you're trying to get at.

MS. BROWN: Just to follow up on

Cathy's comment, I'd like to hear from NCQA about how they think measure 1411 relates to 1390. I mean, yes.

DR. SCHOLLE: 1411 is adolescent well care.

MS. BROWN: I have a feeling that if I knew all of the measures that have been developed and are out in the universe, that there are a lot in this area, and I think that somehow they have to be looked at as a group, and probably come up with one rather than three or four.

I don't know what the absolute number is, and I understand they measure somewhat different things. I do understand that.

But it would be better to have one powerful measure than four, you know, that all have subtle differences. I think it's such a burden on everybody.

DR. SCHOLLE: So there's a challenge to that. So just to elaborate on

what Carol was talking about, so what we have are -- we have this measure that we're talking about, which is about access to a primary care provider. So then we have measures that look at well child visits. And then we have measures that look at emergency department use and inpatient use.

And so we look at -- so we have measures that allow you to look at those different pieces of work. They're based on claims data, and because of the inadequacies of claims data, and the lack of information -- consistent information across health plans or across state Medicaid programs about how they assign or allow, more often allow, patients to identify a primary care physician.

You know, in the claims data, there's some places, some states that could apply this measure to the defined primary care provider for the patient, but we couldn't do that consistently.

So we always come back to the

common denominator of being able to, across the country, so we can have comparative data. So that's the way we've created this group of measures.

And the challenges that you've talked about related to accountability or using this, okay, so those are going to apply consistently across all these different groups.

I mean, adolescents, not all adolescents -- I mean, my kids are healthy, but they have a sports physical every year, right? So that kind of expectation, that's going to be similar across large populations, and we're talking about large populations.

So what you can do is make comparisons, and maybe for age six, or age four, it ought to be 100 percent. For age twelve, maybe 100 percent is not what you're trying to get at.

But this is a utilization measure where we're making comparisons across

populations and over time, so that we can look at trends in access.

And I think that -- we've tried to split it, access, well child visits, emergency use, and patient use. And that gives you a sense of the utilization across, and it would allow us to make national comparisons across states. It allows us now to make national comparisons across health plans.

CO-CHAIR MCINERNY: Just a couple more quick comments and then I think we need to take a vote.

DR. QUIRK: My question is, what's the difference between utilization and access? And I think that they're getting jumbled up and I think that we're kidding ourselves.

I think that all that this measure tells you is how many patients go to their primary care doctor who is a pediatrician. It doesn't tell you anything else.

I think if the health plans want to know about utilization, they can bear the

burden of finding out why patients don't go, because they're the ones paying the premium.

I also believe that access is a very, very complicated thing. Sarah taught me about this 15 years ago.

You know, some of it's affordability, some of it is susceptibility, some of it is availability, some of it is acceptability.

So there's a whole lot of things that go into utilization. This doesn't address that want and form those very, very important questions.

If I'm a working-class stiff with a health plan that is not very rich, and I have to take a half a day off to drive my kid and sit in a waiting room for half a day, my kid isn't going to go.

If I've got a kid who's an adolescent who doesn't play sports, which most of them don't, then I'm not going to take a day off work with no pay to take the kid to

something that the kid doesn't value that to me is a pain in the neck to do.

So there's a lot of those issues. And they're all -- you know, you're in denial. And this doesn't assess accountability. It tells you about utilization at best, but it has nothing to do with access.

I think there is no value to this measure. That's my personal --

DR. GLAUBER: And it doesn't tell you about medical homeness either. It is a measure of utilization.

You know, I realize that this could be used for comparative purposes, but across regions, there may be different models of how primary care is paid and if there is more --

DR. QUIRK: No, because it doesn't -- you're not addressing the issue of co-pays, the cost of co-pays.

Does this contract require a co-

pay? If I go to a doctor, I'm going to incur -- I'm going to get more lab tests, more radiologic and imaging tests, because that's what doctors do. And depending on the structure of the local environment and the plan, I have co-pays. They get expensive.

DR. GLAUBER: Yes, but I'm just making the point that certain systems of care that may be on a more capitated or global budget can facilitate access without promoting visit-based care through other methods of interacting with the system, whether that be Internet-based, advice nurses, and so forth.

So children can have access, and perhaps in more advanced systems of care, without -- you know, to James's point, having to take time off work and churn the kid through the system in order to have a visit for a problem that might otherwise be addressed.

DR. QUIRK: There is no generalizability to this because it is such -

- we are such a cottage industry in the United States. This might work in Boston, all right, but it's not going to work in rural Arkansas.

CO-CHAIR MCINERNY: Okay, we've got to move on. Let's take a vote.

DR. WINKLER: All right. How many of the committee feel that this measure meets the importance criteria? One, two, three, four, five.

How many no's?

Did I have six? Okay, thank you.

One, two, three, four, five, six, seven.

Ellen?

DR. SCHWALENSTOCKER: I would say yes, it needs to be important, this criteria.

DR. WINKLER: Dead tie. All right.

MS. BOSSLEY: I would recommend that you would just keep going through the rest of the criteria and then vote.

DR. WINKLER: I agree. That's

where I was going to go. Glad to hear -- okay.

In terms of scientific acceptability, how many believe it meets the criteria completely?

Partially?

Minimally?

Not at all?

Ellen?

DR. SCHWALENSTOCKER: Minimally.

DR. WINKLER: Okay. Usability. Completely meets the criteria?

Partially meets?

Minimally meets?

Not at all? No?

Ellen?

DR. SCHWALENSTOCKER: Minimally.

DR. WINKLER: Thank you. Feasibility, completely meets? Eleven.

Partially? Two. Are there any others?

Minimally?

Ellen?

DR. SCHWALENSTOCKER: Minimally.

DR. WINKLER: Okay. Are there any minimallies here? She was a min. Okay.

All right. So, recommendation for endorsement, how many vote yes?

One, two, three, four.

How many vote no?

One, two, three, four, five, six, seven, eight, nine --

Ellen?

DR. SCHWALENSTOCKER: I'm no.

DR. WINKLER: Thank you. Okay, it was four yes, ten no.

DR.CHEN: Can I just raise one quick question? And I think that's sort of to echo what Kathy's saying.

So, today I'm getting a little bit more sense of the tension between what's a performance measure and why NQF is endorsing it, and what is it, just a useful quality metrics that any health plan can do on their

own just to keep track of whatever they're doing.

So this is obviously not a perfect measure, but it is a measure that I think can provide useful information. How useful it is depends on how creative the person who is analyzing it -- is doing with it, but it is a useful matrix, nonetheless.

But obviously, it's not up to the standards of a performance measure, especially for high-stakes performance measure.

So my question is, maybe this was asked yesterday as well, but I'm a little bit more confused today than I was yesterday. Does NQF endorsement lead to support and funding for NCQA to do this more, to have more elaborate testing, or field testing? Or is it just give them a ground where health plans have to try to look at these numbers? Or, I mean, what's the point?

DR. WINKLER: NQF is looking to evaluate measures and endorse them for the

major purposes for public reporting, as well as of quality improvement.

Throughout the decade of NQF's existence, we've certainly seen a focus of using NQF-endorsed measures when measures are being used for public reporting.

We've certainly seen that by the federal government, and we see it with a lot of private plans and others. And so that is what NQF is looking -- that's our goal, is to, of all the measures out there, to sort through them and find the ones that meet our criteria such that they can be used for those purposes.

DR. CHEN:: I'm sorry. So then I don't understand why this measure couldn't be used for public reporting.

I think Cathy's concern is that it may be used for pay for performance which may not be fair, but for public reporting a cross-comparison of health plans across the nation and following trends over time, I don't see any issue with this measure being used, as

long as we know its limitations.

We know it's actually an utilization measure. It's not an access measure, I guess, if you want to make that distinction.

DR. WINKLER: One of the other things I'll just mention as a result of NQF's processes, we deliberately bring different perspectives to the table.

You all represent sort of different points of view, and the measure evaluation criteria are attempting to define a type of measure. And there has been a deliberate attempt to raise that level of the information and the quality of information that you get over time.

And so what I could say is a result is not everybody at the table agreed with you.

DR. LIEBERTHAL: While the NQF's goal is for public reporting, in fact, the insurance companies are looking to NQF-

endorsed measures for their pay for performance programs.

And I would have to make the assumption with each one of these measures that it may be a pay for performance -- used for pay for performance. So I don't think we can separate the two.

DR. JENKINS: That's explicit, in the state of Massachusetts, they're required by statute to only include endorsed measures for pay for performance.

MS. BERGREN: So the plan level criteria that we look at could be used for pay for performance?

DR. LIEBERTHAL: The purchasers of insurance from a plan such as, in our occasion, California, the Pacific PGBH, it's a big consortium of companies that have gotten together to evaluate health insurers and health plans. And they do use these types of measures in determining which insurers they will contract with and how much they will pay.

DR. JENKINS: And also some of the plans, because they need to meet these specs, roll it down further within to the practice level as an accountability measure. And that's where that bright line starts to come from.

I think, Alex says a population health measures, I would have endorsed it. It's that accountability at the plan level that I'm struggling with.

DR. WINKLER: We have another measure developer waiting patiently to talk about the next measure, and I think maybe after that we can take a break.

The next measure we're going to talk about is measure 1419. This is primary care prevention, intervention as part of as well or ill child care as offered by the primary care medical providers.

This is from the University of Minnesota, and Dr. Rao?

DR. RAO: Yes, I'll just give a

brief introduction because I know Dr. Diener's on the phone.

So this is a revised measure, essentially. Dr. Diener was here last year to present the measure, and it deals with the application to fluoride varnish to children at risk for caries, and that includes the Medicaid population. As he points out that that's only 30 percent of children, but they represent about 80 percent of the disease burden.

Furthermore, it's an important measure, because the fluoride varnish has been shown to reduce the incidents of caries significantly.

There is a great deal of room for improvement because this is just emerging. Funding through the EPSDT programs is now available in about 40 states, and seems to be becoming much more common.

So the measure is, the extent to which the fluoride varnish is applied during

EPSDT visits among Medicaid and high-risk children.

The only -- in terms of its usability, feasibility, I think it's relatively easy to use because it comes from claims data.

It's likely to be accurately reported, because it represents a supplemental payment on top of the EPSDT visit. So my practice manager, my practice has said, you know, if you're doing this, make sure that you mark this off, because we get $18 or $19 extra, so it's something that we can probably measure fairly easily here.

And I think it's a very progressive measure, because it's looking at a physician-based or a practice-based behavior that's relatively new, whereas a lot of the things we've talked about already, we should have been doing for the last 50 years, and we were just not doing them right.

So, I think it's progressive in

that sense. It can lead practice a little bit.

I didn't see -- and Dr. Diener can clarify a lot of information about testing. I know that the measure is in use right now, but in terms of results and things, I didn't find those in the measure application that's here.

So I'm going to maybe stop there and let him provide some more information.

DR. DIENER: Thank you very much. I'm a little unclear as to what it is that you're looking for that you didn't find.

DR. RAO: Just results on how well the measure is performing, how well primary care providers in Minnesota, for example, are doing with it.

DR. DIENER: Well, I can tell you that between -- based on -- this is all based on claims data and a report generated by the Department of Human Services that between 2008 and 2009 the number of EPSDT exams with

fluoride varnishing included rose across the state now from 5 percent of EPSDT visits to 8 percent of EPSDT visits.

It's measurable, and if you track over time, granted, it takes a year -- DHS considers that the data set is not complete until a year out from the last date of service in the year. So it's not going to be until December 31st of this year that the 2009 data set will be considered complete.

So I'm giving you a snapshot along the way. So, it can be measured. It can be teased apart by whoever's provider number is used for the billing.

Oftentimes, it's a physician. Sometimes it's a clinic. Sometimes it's a health plan.

And if you introduce the patient ID number you can actually tease out from this data set the number of varnishings each individual child got during the year from anywhere from one to x number of providers.

They may switch clinics four times during the year and get varnished from four different providers during the course of the year.

DR. PERSAUD: I have a question. Does this exclude varnish done by a dental provider?

DR. DIENER: Absolutely.

DR. PERSAUD: So this is limited --

DR. DIENER: All medical claims.

DR. PERSAUD: So, okay, so this is on medical claims data.

DR. DIENER: Yes.

DR. PERSAUD: And I wasn't clear on the numerator and denominator regarding whether you're assessing number of applications per age of the child against expected number of applications, or just the raw number. Is it a rate? Is it a percentage versus expected?

DR. DIENER: Now, at the most basic level, I'm looking at it statewide as a

measure of how well physicians or primary care medical providers, because we've got to include physician assistants and nurse practitioners, at least in Minnesota, how well primary care medical providers are providing carriage prevention services to their patients during the course of an EPSDT exam.

DR. PERSAUD: So is it measured against the expected number of applications they should have had for their age? Is it a percentage of the expected number, or just how many applications?

DR. DIENER: Well, you can -- if you start getting into individual patients by using the patient ID number, you can break it down during the course of each year, how many did the child get?

Now, the child, according to the American Dental Association's recommendations, starting with the eruption of the first tooth or by age one, every child -- high-risk child should get four applications a year, all the

way through teenage years.

So the number four is the ideal for the year for child regardless of age. You can break it down if you introduce the patient ID, how many individual patients got it. But at the gross level, I'm looking at the number of EPSDT exams Dr. X did and how many varnishings he did as part of that number of EPSDT exams.

Ideally, it should be 100 EPSDT exams and 100 varnishings. And you've got every kid at every EPSDT visit.

MS. CARLSON: Does Minnesota not provide a dental benefit? Is that why --

DR. DIENER: Oh, no.

MS. CARLSON: -- you're only looking at medical?

DR. DIENER: No. We're looking at medical across the country because dentists generally won't take Medicaid children. That's the problem.

MS. CARLSON: Right. So, but in

those states where dentists do take Medicaid children, is your expectation that physicians still reach 100 percent of the EPSDT screens?

DR. DIENER: Nationwide, dentists only get paid twice a year for putting varnish on. So, in the best of all worlds, a Medicaid child has a dental home, i.e., a dentist who will see the child whenever the child has a problem and regardless of what the problem is, the dentist sees the child twice a year and puts varnish on.

The dentist communicates that care to the physician, who then on two visits for medical care during the course of the year, three months apart from the dental visit, puts varnish on.

Now, the child gets four varnishings a year, twice at the dentist, twice at the physician's office. This is the ideal.

We're a long way from ideal because so few kids get in to see the dentist

and in greater -- you name the state, Minnesota, Massachusetts, Kentucky, it doesn't matter.

In grading, in the rural parts of the state where there are more general dentists than pediatric dentists, general dentists are scared to death of one and two year olds. They've had essentially no exposure to them in dental school. So they will tell the mothers, starting at age three, when the mother has been told by her public health nurse or her physician to start at age one.

DR. RAO: Dr. Diener, just a quick question. I mean, you're saying that there's no EPSDT exam where this varnish is not indicated?

DR. DIENER: Well --

DR. PERSAUD: That's not true for the state of Texas.

DR. DIENER: Well, let's see. In the first year of life, we see kids at two

months, four months, six months, nine months.

Okay. The child won't start getting teeth until six months or thereafter. So in theory, you might have two EPSDT exams in less time than the three months' interval that the ADA says should be observed for putting varnish on.

But once you hit age one and you see kids maybe every six months, you may end up having to have special fluoride varnish clinics akin to the immunization clinics if you really want to work hard at getting the kid varnished four times a year.

DR. PERSAUD: I guess what I'm struggling with is just trying to look at, for example, the state of Texas, where Medicaid has decided to pay for this up to the age of 36 months only, and to allow, I think five or six total. That's the situation.

And then my practice is in an urban environment where we actually have lots of dentists that take Medicaid. So, and we've

trained, and can and will do this, and I think the children will get a blend of both types -- both applications.

And I'm just kind of wondering whether this should be all on the medical provider, or let's get all the codes, because what they should get is the dental varnish applied, just so that we don't end up with a rigid construct that does not flex to the different practice patterns.

I think, to me, the bottom line is, the oral varnish needs to get on those teeth, no matter whether it's a dentist or a medical home.

DR. DIENER: I agree with you. And ideally, if we, doctors and dentists, bill on a similar billing system, you could merge the two.

But dentists bill on an ADA form that is quite -- and it goes to the dental insurer within the state versus the medical side.

I'm not sure that the state has the capability of bringing together whatever claims data they may get from the dental side together with the claims data they're getting from the medical side.

At least on the medical side, they've got to roll things up for the CMS 416 report. And in Minnesota at least, the report I referred to earlier is an offshoot of the report that has to be created for the CMS report.

And Minnesota's got one person at DHS who can do reports. That's the extent of their staff. So there's a staff limitation. I think, at least in Minnesota, it costs money.

So, ideally, yes, it would be wonderful to merge the dental and the medical claims database. Then you could find out to the last varnish how many varnishings a kid got during the course of the year from both the doc and the dentist.

DR. GLAUBER: And I'd just like to underscore Goutham's point about this being a progressive measure, especially viewed with respect to the clinically preventable burden of disease here.

The target population is very similar to the target population for the lead screening measure, which is a HEDIS measure, and most Medicaid agencies hold plans accountable for, and therefore, resources and improvement effort is directed towards that particular measure, which doesn't have U.S. Preventative Services Task Force endorsement.

So if the committee thinks that there's really an opportunity here to prevent -- to make a dent against one of the most common conditions in childhood, especially in the Medicaid population, I think having a measure here would really spur improvement efforts.

DR. DIENER: I'm sorry, I missed the last couple of words of your sentence.

Will do what?

DR. GLAUBER: Well, I think having an endorsed measure here will focus greater intention within state Medicaid agencies and plans towards improvement efforts.

DR. JENKINS: Are you suggesting the opposite to everything I've said at all the other measures, that there should be a plan accountability and that they do have access to both the dental claims and the medical claims?

DR. DIENER: Ma'am, excuse me for a second. As long as I've been on the phone call, yours is the only voice I cannot understand on the phone. It's just -- I can't pick up any words. There's something about the microphone that you're using is different from all of the other microphones in the room.

DR. JENKINS: Yes, I agree with everyone. I thought this was one of the most important measures we saw under the outcomes work, and now here we are again under the

process measures work, and I'm hoping we can find a way to solve the measurement issues.

I guess my question is to Dr. Glauber, where I personally have argued against plan accountability on many measures, perhaps this is one where we could hold the plans accountable and they would solve the measurement problem by having access to both the dental claims data and the medical claims data in a way that no one else will.

DR. LIEBERTHAL: The denominator is a very complex statement that implies that a risk assessment be done without stating what the risk assessment tool is. It includes primarily Medicaid and CHIP- eligible children, but I think it equally applies to commercial insurance children.

The communication with the dental home, under the best of circumstances, is minimal, if any. You know, the best I can do is ask the family, have they seen a dentist for the child, and I assume that the dentist

that they saw now becomes the dental home.

DR. DIENER: Well, I'm not sure that that's a fair assumption. A lot of kids will see a kid once to pull a tooth and then that's it.

DR. LIEBERTHAL: Well, it's pretty hard then, for us, to determine whether a dental home -- the child has a dental home. So, as you look at the denominator, I don't think it's a functional denominator.

DR. DIENER: The denominator is all children who have a child and teen checkup.

DR. LIEBERTHAL: Well, that's not what the denominator statement as submitted says.

It's a long paragraph, all high risk children, dot, dot, dot, will be identified by paper and pencil caries risk assessment tool. And if the child is covered by Medicaid, CHIP, but does not have a dental home and then the child is high risk, if a

child is, but they do have a dental home, other risk factors will be considered. A very complicated statement to implement.

DR. DIENER: Well, in fact, there are two questions that are asked first, are you on Medicaid? Yes. Do you have a dentist who will see you regularly? No. You're high-risk.

Basically, most of these kids who are on Medicaid or CHIP, by virtue of their inability to get regular comprehensive dental care, are high risk.

But the ADA, in its fluoride varnish recommendation, distinguishes between low risk, moderate risk, and high risk.

CO-CHAIR MCINERNY: But that's not what the denominator statement says.

MS. BROWN: Just two questions. What does the AAP think about this? I mean, this is now saying to pediatricians, and other primary care providers, that they need to take this on. I'm just wondering what their

position is.

And then secondly --

DR. DIENER: Well, I'm part of the AAP Oral Health Initiative. We're going around the country trying to get the seven states that are currently not yet reimbursing to reimburse, and we're working across all the states where reimbursement is in place to get as many primary care pediatricians to do fluoride varnishing as part of well child care.

MS. BROWN: Okay. And second question is, Carol and I were just muttering, in essence, this lets the dentist off the hook.

It says, I don't know why you're not doing this, we wish you would, but we're going to fill in the gap. And I understand that.

But what's the long-term strategy to get dentists to respond to this need, rather than just say, you're not doing it, so

somebody else is going to fill the void, or is that not --

DR. DIENER: Well, some of that's going to have to come from within the dental profession, and I can't speak to that.

But I am going very soon now to start a project in Minnesota with the State Health Department where we're going to try to create in greater Minnesota oral health zones, getting all the stakeholders together in a community, tell them the extent of the prevalence of caries in their population of their children, and trying to get them to take ownership of their problem in their community.

Ideally, then, the physicians and the dentists in that community will figure out how to work together. Ideally, the physicians can do the prevention and the counseling of the caregiver and putting on the varnish if the dentists will take the train wrecks who need restorative care.

But there's no guarantee that

that's going to fire. I mean, the dentist has to be willing to participate.

CO-CHAIR MCINERNY: This is Tom McInerny. And I was present for this measure when it was looked at previously here, and we did get hung up on the denominator, and I think we're getting hung up again.

Now, the original presentation a year or so ago, you used the term, EPSDT, but you now have substituted a term -

DR. DIENER: Well, in Minnesota, the EPSDT exam is called child and teen checkup.

CO-CHAIR MCINERNY: Child and what?

DR. DIENER: Child and teen checkup.

CO-CHAIR MCINERNY: Oh, child --

DR. DIENER: That's Minnesota's name for the EPSDT exam.

CO-CHAIR MCINERNY: All right. Well, thank you. But I still think your

denominator statement is far too complicated and needs to be simplified significantly.

DR. DIENER: Well, I was advised that what I submitted had addressed the issue. I'm sorry that -- it is what it is, but I was told that everything looked great. I mean, I can redo it and set it up, the denominator, all kids who get a chart are on the EPSDT exam, period.

MS. BOSSLEY: This is Heidi. I mean, if this denominator said it was all exams, would that -- I mean, it sounds like that might address the committee's concerns. It sounds like you might be able to do that.

DR. DIENER: Right, because the report that I get from DHS shows by provider the number of duplicated and unduplicated EPSDT exam done during the course of the prior year.

The other two columns state the number of fluoride varnishings done duplicated and unduplicated.

DR. RAO: You know, I just want to add that that would solve a lot of problems, because if this is a progressive measure, what we're monitoring for is improvement over time.

So if they have a dental home, if they have -- if they're getting care elsewhere, even if they're not high risk, we'd still expect each provider's performance to improve, so that would be helpful.

DR. DIENER: And really in very simple terms, the number of kids who should get an EPSDT exam and the number of kids who as part of that EPSDT exam got fluoride varnish. And you can parse it by child after that.

DR. PERSAUD: The other issue is the one about age, that this goes up to age 20. And I doubt that there is uniformity amongst the states for paying for this up to the age of 20.

DR. DIENER: No. Minnesota is, and a couple of other states are, but you're

right. Some states pay up to age three or age five or age six or age thirteen.

Each state Medicaid program has a call on the number of varnishings it will pay for per year, the age of the child who gets the varnishings, the training that the provider has to undergo, the codes to be used. All of that is at the state level.

DR. PERSAUD: So I guess the question for us is, you know, what will this measure drive, if it's not reimbursed to the same age in different places?

Do we expect it to drive reimbursement, and then to realize from an accountability perspective, you would get into issues if you're practicing in the state where Medicaid is not paying for that benefit past a certain age.

DR. DIENER: My hope is that as time goes on, those states which are reimbursing only up to age five or age six will begin to look at some numbers and see

that they are saving money on expensive ambulatory surgical care restoring multiple carious teeth in ambulatory surgery with the risk of anaesthetic death, ER visits for abscess teeth, which is only incomplete care because ER doctors don't pull teeth. They drill and fill, they tell you to see your dentist in the morning, and then they treat pain and infection.

I'm hoping that the states, unlike Minnesota, which is already paying up to the age of Medicaid eligibility, will move in that direction.

There's no guarantee, but at least, at least every state that is paying, and there are only seven that are not, Hawaii, Oklahoma, Louisiana, Arkansas, Tennessee, Indiana, Delaware, and the District of Columbia, are the only states that are not at this time reimbursing. They're all reimbursing up to age three.

DR. LIEBERTHAL: Two points. One,

I think as we determine what states are currently paying should not be an issue if it's scientifically valid and a quality improvement measure, then what states are currently paying for it should not be an issue.

Second, when we vote on this, can we vote on it based on a simplified denominator?

DR. DIENER: I'm perfectly comfortable with that.

DR. PERSAUD: So can someone summarize what that denominator is?

DR. LIEBERTHAL: As I interpret it, it is that children at EPSDT visits will get the varnish.

CO-CHAIR MCINERNY: Right. That's it. The denominator is all children with an EPSDT exam.

MS. BROWN: I've got two data questions, though. If the dental data and the primary care practice data in Medicaid can't

be blended or put together in some way, that's a real problem to this recommendation.

If there's a child who is getting it at a dentist but whoever brings that child in for an EPSDT visit doesn't know that, so I think that we need to understand whether or not this recommendation hinges on the ability to link those two data sets.

And then secondly, if a state Medicaid program isn't going to pay for varnishing, then saying that it's to be done up to age 20 is not useful.

I mean, sort of as a political organizing effort, it is, but in essence, then, for a state that doesn't pay for it, then you're asking the dentist or the primary care provider to do this for free. And I don't understand that.

CO-CHAIR MCINERNY: Yes. It's what we call an unfunded mandate.

DR. SCHWALENSTOCKER: This is Ellen. With that thought in mind, I'm

wondering if -- it sounds like this measure is being considered as a provider or a plan-level measure, maybe.

It might be helpful to make it a population-based measure, initially, and not hold individual --

DR. JENKINS: But then be holding Medicaid responsible, not the practitioners --

DR. DIENER: I can't understand what you're saying.

DR. PERSAUD: For a cleaner measure, I'd feel better if the measure went to age five and it was any oral varnishing, dental or medical. Because that's what I'm interested in, are the children getting the varnish, or not? And if they're not, etcetera, I mean, that's all in the measure.

But that would be a cleaner measure that is interpretable. Then we can understand that if they're not getting it, what it does mean, perhaps, is that the dentist just can't come through, and then it

may go all medical.

And then if we go up to age five, and that's an age -- or age three, where the threshold number of states do cover it, well, then at least we can understand that there may be barriers beyond just reimbursement and deal with those and move on.

CO-CHAIR MCINERNY: I agree. I think it's time that we take the vote here. We're running on. So, let's go ahead.

DR. MILLER: Tom?

DR. MILLER: Tom, this is Marlene Miller. I have a question. When you're going to take a vote on this, it sounds like we're talking about changing the definition.

And if the committee wants you to change a definition, all the prior reliability and validity testing is out the water, so I don't understand exactly how we're going to do this process.

DR. WINKLER: Well, I think that's the question at hand, and maybe as an

alternative to voting, perhaps we need to, as we've done with some of the other measures, go back and kind of put the words in place so that everybody knows what you'd be evaluating.

But it sounds like the developer is sort of amendable to a lot of this stuff, so I think if we make it a little bit cleaner, and then the committee knows exactly what they're voting on.

DR. MILLER: But then all the testing that's done prior to now doesn't matter, because it's a new measure.

DR. WINKLER: I think that some of the questions that the committee was asking actually were clarifications. And I'm not sure that -- I think we need to find out if it really changed anything or more just make the meaning more clear.

Does that seem -

MS. BOSSLEY: Reva, this one would be for time-limited anyway.

DR. WINKLER: Okay.

MS. BOSSLEY: So there wasn't any testing provided.

DR. WINKLER: All right. But I still think there's enough -- we need to clarify the words for everybody.

MS. BOSSLEY: Yes, I agree. I think you all could probably table this measure, wait until you get something revised back, and I think, again, it would be time-limited.

DR. WINKLER: Right.

MS. BOSSLEY: That doesn't change.

DR. DIENER: What does time-limited mean?

MS. BOSSLEY:: You would have to provide testing information on reliability and validity within 12 months of endorsement. But we can talk more offline on what that means for you.

DR. MILLER: This is Marlene again. I thought I was at a meeting with Helen Burstin and she said we're not doing

time-limited endorsements anymore.

DR. WINKLER: You're breaking up a lot, and we're not really understanding what you're saying very readily. Are you on speaker phone?

DR. MILLER: Yes, I've picked up the handset. Can you hear me better now?

DR. DIENER: Much better.

DR. WINKLER: Yes, that's better.

DR. MILLER: I thought I was at a meeting with Helen Burstin about a month ago and I thought time-limited endorsements, we're not doing anymore.

DR. WINKLER: There are limited times when we do them, and this particular project, because there is an association with the CHIPRA efforts around the set of measures, that is one of the sort of exceptions. And so for this particular project, that option is still open.

MS. BOSSLEY: Right. And Marlene, just to give you a little bit more detail on

that, there's really three components to it.

One is a legislated mandate or some expected need for measures. The other one is there's no other measure within this category or topic and this measure really does apply. And I'm, as usual, blanking on the third one.

But if it's within that -- oh, it's not complex. So it's not a composite or outcome measure.

DR. WINKLER: All right. So, we agree we'll table it until we can get things a little more clarified, and then bring it back to you for a final decision?

Break time?

CO-CHAIR MCINERNY: Break time. Ten minutes. Twenty of.

(Whereupon, the above-entitled matter went off the record at 10:29 a.m. and resumed at 10:44 a.m.)

DR. WINKLER: Okay. We're going to continue on the topic of oral health,

dental visits, and we're going to start with measure 1405, oral health access, from NCQA.

Martha?

MS. BERGREN: Okay. This is a process measure at the primary care provider level, and the priority is population health and care coordination.

And the oral health access is part of the composite well child care at ages two, six, thirteen, and eighteen. And this is a four-measure criteria at each of those age ranges, and it's a chart review criteria.

Again, this is oral health access, which is the most common childhood chronic condition. And is also a preventable issue. And when damage occurs, it is irreversible, and this was all well documented.

Some of the concerns from the members on the importance was, who is the one responsible when there is non-compliance on this in the consumer's choice of whether or not they access dental care?

And that there's some disagreements about whose job this is, the primary care provider's or the dentist's responsibility.

So the groups that are endorsing this are the American Academy of Pediatric Dentists, the ADA, the AAP and its Bright Futures measure.

There's also the United States Preventative Services Task Force recommendation for the oral fluoride supplementation, which is a B recommendation. But there's insufficient evidence to recommend for or against routine risk assessment of preschool children by primary care clinicians for dental disease prevention.

There's no validated risk assessment tools, little evidence that the primary care provider can systematically assess for risk, little evidence that counseling or referring high-risk children leads to fewer caries or reduced dental

disease, and that referral does not result in subsequent dental visits.

The numerator is that an oral health screen is documented in the medical record at ages two, six, thirteen, and eighteen, and the note must mention at least one of the following: dental treatment by the PCP, risk assessment by the PCP, referrals to the dentist, or parent report of a dental visit.

The denominator is that the child who has turned that age within the measurement year and had a face-to-face visit that predated the child's birthday by at least by 12 months.

No risk adjustment, no reliability testing, the validity is the panel of experts face validity. And again, the method of extraction is the chart review. And they estimate that a random sample of 30 to 50 patients in a 2,000 patient practice would be sufficient.

The measure is not currently in use, but it was field tested. The usability was tested, and the feasibility is the chart review. It's difficult.

That's it.

CO-CHAIR MCINERNY: Question. The way this is phrased, it's the responsibility of the PCP to inquire at least about dental visits and/or do risk assessment. Is that correct? Is that the limit of the PCP's responsibility?

MS. BERGREN: To ask and either assess risk, refer to a dentist, or document that the patient reported that they have had a visit.

CO-CHAIR MCINERNY: Refer to a dentist with a note saying that the PCP advised the parent that the child should be seen by a dentist, is that considered --

MS. BERGREN: It didn't go into that detail.

DR. SCHOLLE: It also includes

treatment, so the kinds of treatment that were addressed in the previous measure, so sealants or the varnishes, any of that treatment that occurs would also count, and a referral would be that the referral is documented in the chart.

MS. BERGREN: The group that --

DR. SCHOLLE: It's a specific referral, not just a referral generally, you should see a dentist, but a specific referral.

MS. BERGREN: The group that reviewed this felt 100 percent that it was important and that there was some differences of opinion on the outcome measurement, but not a lot of concerns about this measure.

DR.CHEN: Just a question for Sarah. Referral, you meant actual referral in terms of insurance referral to like a dentist, or is it -- I think you said that just writing down that the patient is referred to a dentist is not enough, right? It's not sufficient?

DR. SCHOLLE: So, different

insurance companies would have different rules about how to get it, but what we wanted is a specific referral to a dentist.

DR.CHEN: What do you mean by that? I mean, is it a referral form, an appointment made, or --

DR. SCHOLLE: So, we didn't specify that an appointment needed to be made, but that there was a specific referral to a dentist that was documented.

DR. QUIRK: I need clarification. Being someone who sees patients, when you say that a referral be made, you're saying that I have onus of making the phone call and finding a dentist who will -

DR. SCHOLLE: A patient referral to a dentist. So it wouldn't -- I mean, it's --

DR. QUIRK: I could make a recommendation to the patient. I refer to a dentist.

Okay, just for precision, in terms

of the terminology as it is generally used in health care. Do you understand my -- because to have to call and make a referral puts an enormous burden on a practice.

DR.CHEN: So, just to give an example, where I am, we don't actually have dentists -- our dentists no longer take our kids, basically, because of insurance reasons.

So what we do is we generate a list of community dental providers that would see these kids, and we give them that list. Would that count as referral?

DR. SCHOLLE: Yes.

DR. QUIRK: Okay.

DR. GLAUBER: Question for NCQA. Since this is a health plan measure and requires chart review, this would be based on a sample of children, correct? So --

DR. SCHOLLE: This is actually specified at the provider level rather than at the health plan level.

MS. BERGREN: We have a different

issue, that it's the HEDIS measure, but this is the physician level.

DR. SCHOLLE: So this is intended to come from the chart, but so the entire set of composite measures that -- when we've said that it's part of a composite and it's provider-level, those are based on reporting from a provider organization, a practice, or a physician.

DR. GLAUBER: But are they pulling a sample of kids who have had a face-to-face visit in the last year, or a comprehensive well visit?

Because I could only imagine that this level of activity occurs during a well visit. So if we're also including kids that didn't have a well visit but just had an illness visit, then that's going to contribute to variability in performance.

DR. SCHOLLE: So in the denominator for these provider-level measures, we struggle with, how do we define the

population that that provider is responsible for.

In the health plan, it's people that are enrolled, and usually we have some kind of continuous enrollment criterion.

For provider-level measures, the way we've managed that is we look for somebody, for the child who's had their second birthday and who have had evidence of an ongoing relationship with this provider as determined by they had a visit that predates the child's birthday by at least 12 months.

So, what we're saying, they're selected -- the way the sampling selection goes is that they take a date and they take a consecutive sample of children who have had a visit.

And then we look to see, has this provider seen this child -- seen this child sometime before a year ago, because that tells us, well, at least that child was having care from this practice for a year?

Now, I can see that sometimes kids would go and come and go and come, but generally, what we're trying to do is establish that there's been an ongoing relationship with a practice, and so a visit -- but not necessarily a visit during that measurement year during that 12-month period.

And all of these measures have a look-back period of two years, except for -- so that if they addressed it eighteen months ago, it would still count.

We don't expect them to have a visit -- and not all children of the age groups that we're looking at are going to have a visit every year or a well child visit every year, so we're trying to allow for that.

DR. WINKLER: Any further discussion? Are you ready to --

DR. JENKINS: Could I ask what, when it says a risk assessment performed by the primary care clinician, what would a risk assessment be?

DR. SCHOLLE: I'm looking to see if our specs have details on that. We don't have specific language, other than I believe that we were building that term based on our knowledge that AAP was recommending this risk assessment.

So what we wanted to see is that they were looking -- we originally tested this with something that said, did they look in their mouth, and that wasn't good enough.

And when we got back from our panel, you know, we said, okay, did they look in their mouth? Did they do other stuff?

So what we wanted to do was get some sense of whether they were thinking about dental risk and so we left it at that. This is --

DR. WINKLER: Decision time? Okay.

How many on the committee feel this measure 1405 meets the importance criteria? Thirteen.

How many say no?

And Ellen and Marlene, you're on the phone?

DR. SCHWALENSTOCKER: I say yes.

DR. MILLER: I agree.

DR. WINKLER: Both yeses, thank you.

Okay. So moving on to the scientific acceptability of the measure, how many feel the measure meets the criteria completely?

Partially? Ten.

Minimally?

Not at all?

Okay. Marlene and Ellen?

DR. MILLER: I'd say minimally.

DR. WINKLER: Okay. Ellen?

DR. SCHWALENSTOCKER: I vote partially.

DR. WINKLER: Okay. And for usability, how many feel it meets the criteria completely? Two.

Partially? Four.

Minimally?

Are there any not at alls?

Okay. And Marlene and Ellen?

DR. MILLER: Minimally.

DR. WINKLER: Ellen?

DR. SCHWALENSTOCKER: Partially.

DR. WINKLER: Okay. Feasibility, how many feel it meets it completely? None.

Partially? Seven.

Minimally?

Any nones?

Okay. Marlene and Ellen?

DR. MILLER: Minimally.

DR. WINKLER: Okay. Ellen?

DR. SCHWALENSTOCKER: Partially. I guess I'm a higher grader.

DR. WINKLER: Okay. Not a problem. And in terms of recommendation for endorsement, how many would vote yes? Four. Okay, okay.

How many would say no?

Marlene and Ellen?

DR. MILLER: I'm a no.

DR. WINKLER: Okay. Ellen?

I can't hear you.

DR. SCHWALENSTOCKER: All right, yes.

DR. WINKLER: Again, another close one. It was seven yes and eight no. Yes. Okay.

All right. We have another dental visit measure, 1388, so we'll move into --

MS. BERGREN: Okay. 1388 is the percent of members two to twenty-one who had a dental visit in the calendar year.

It's an access measure at the health plan level, and a population health care coordination measure, and will be collected via claims data, administrative data.

Again, all the importance was documented. It's the most common childhood chronic condition. It's irreversible. Once it occurs, damage occurs.

Again there was concerns about who is responsible for the outcome. Recommended by the American Academy of Pediatric Dentists, the ADA and the AAP.

And again, the same concerns with the Preventative Task Force recommendations.

The numerator is one visit in one year detected on the claims encounter, and the denominator is all plan members ages two to twenty-one.

There's no risk adjustment, no reliability testing. There's face validity, and it's a current HEDIS measure. And there are not multiple data sources.

There has been some testing done. The feasibility, there were some concerns about the feasibility. When dental visits are not covered by insurance, this might lead to some inaccuracy in the medical claims data.

And 23 percent of children don't have dental insurance. And 29 percent of children that do have dental insurance have

Medicaid, and they are much less likely to be able to access a dental care provider even if they have insurance.

Okay. That's all I've got.

DR. LIEBERTHAL: I don't see how a health plan can be held accountable for this, because in our bifurcated world of medical insurance and dental insurance or very often non-insurance, the health plan has absolutely no control over whether a child sees a dentist.

All we can do is recommend that they see a dentist, but whether they actually see one is far beyond the control of the medical health plan.

DR. SCHOLLE: This measure is only for Medicaid health plans and for children who have a dental benefit as part of their Medicaid coverage, so that's covered in the specifications.

DR. LIEBERTHAL: And, but again, the health plan would not be -- have control

over the -- whether the child exercised that insurance, and they're two entirely different -- well, I don't know if they're different billing systems, but the health plan doesn't have access to that dental billing system.

MS. CARLSON: I guess speaking as a health plan in a state where Medicaid does offer the dental benefit, health plans do have access to the dental claims, and they are held accountable.

Actually, it's a P for P measure. It has been in place for several years now, and it seems to be working, although, so access to data shouldn't be the issue.

I think the greater issue is access to providers, which is sort of a universal issue nationwide. So, I guess as a health plan, I wouldn't consider this an issue in terms of data retrieval or extraction.

DR. GLAUBER: And I think there's a whole range of HEDIS measures for which health plans are held accountable where we

would argue that we have limited ability to impact the outcome.

You know, colonoscopy screening, there's only so much your health plan can do to get you to do that, but yet we're measured by it, and we certainly undertake effort to encourage and educate members about the importance of it.

So, I don't think that should be a disqualifier for a measure that a health plan has limited ability to impact the outcome.

DR. LIEBERTHAL: Well, for something in the medical field, the health plan can outreach to the patient and arrange for the visit with the specialist who does the procedure, you know.

And they do have control, mammography, colonoscopy, pap screening, immunizations, of course. But when you cross over -- at least in my view, when you cross over to the dental field, the ability to impact it in any way lapses. I may be wrong.

I may be wrong.

DR. SCHOLLE: May I just clarify? The specifications take that into account. So this is -- for Medicaid plans, this is specified as a measure for Medicaid plans only because dental is often -- it's an allowable benefit under Medicaid. Not all states exercise that optional benefit.

So, however, in the specifications, we go further and say the measure only applies to children who have the dental benefit.

So, in that way, we've tried to take into account those concerns about whether or not the health plan -- whether or not this is a covered benefit for this particular child, so that when we're making comparisons across health plans, we're looking at health plans that have this as a benefit for their children.

And it does -- it's really an access to dental care measure. So it's

looking to see whether they had a dental visit, and it can identify problems.

We realize the problems in networks, and availability of pediatric dentists or dentists who take Medicaid. The issue is to be able to make comparisons across states or across health plans.

DR. LIEBERTHAL: Yes, I'm not arguing about the importance, or that this is an extremely important issue, just where the accountability lies.

And I'm speaking from the perspective of California and an HMO in California, where I think it would be very difficult.

Now, again, I don't know how other states work, but in California, which is the MediCAL, which is our Medicaid program, contracts with companies that have been formed for the express purpose of following the care of Medicaid patients.

And I could see -- so in our area

it's called LA Care, so I could see LA Care being held accountable, because they're totally accountable for all the care the child receives.

But I couldn't see Kaiser being held accountable. So I think it's just a matter of where you designate accountability.

MS. CARLSON: And I think for this measure, it is designated down to the health plan level. If the health plan has contracted with the state for the Medicaid dental benefit, then this measurement would apply to that health plan.

If the health plan has not contracted with the state to provide that benefit, then they would not be measured using this metric.

DR. ZIMA: I think this additional information is really helpful, and I would recommend that some of these caveats, even though it's stated in here, it says HEDIS.

For those that may not be HEDIS

gurus, if it could be repetitive and built into that numerator, I think it would help.

DR. PERSAUD: I would say this is probably a progressive measure in that this is something that can't be substituted for in children's care.

And I was asked, you know, well, isn't oral varnishing enough? And no, it's not.

They do need dental care, and the primary care pediatricians have had multiple measures in place. And the plan I have seen work with providers who aren't doing checkups.

And I think maybe it's an opportunity for plans to work with dentists, because we don't have a replacement for this yet.

DR. WINKLER: Okay. Decision time. For the committee, for this measure, 1388, annual dental visits, how many believe it meets the importance criteria?

Yes?

Marlene and Ellen?

DR. MILLER: Yes.

DR. WINKLER: Ellen?

DR. SCHWALENSTOCKER: Sorry, I was on mute. Yes.

DR. WINKLER: Thank you. Are there any no votes? No, okay.

All right. Under scientific acceptability, how many believe it meets the criteria completely?

CO-CHAIR MCINERNY: Can I ask, is that if we're understanding that this applies only to those plans that provide a Medicaid dental benefit?

DR. WINKLER: Okay. Yes. I think we will ask NCQA to maybe make that a little bit more explicit in the specs, to make it very clear for everyone.

Completely meet the criteria for scientific acceptability?

Partially meet? Six.

Minimally meet?

Ellen and Marlene?

DR. MILLER: Partially.

DR. SCHWALENSTOCKER: Partially.

DR. WINKLER: Okay. And for usability, completely meets, how many?

No, partially meets?

Minimally? Okay.

Ellen and Marlene?

DR. MILLER: Minimally.

DR. SCHWALENSTOCKER: Partially.

DR. WINKLER: Okay. And feasibility, completely meets?

Partially meets?

Minimally? One.

Marlene and Ellen?

DR. MILLER: Partially.

DR. SCHWALENSTOCKER: Partially.

DR. WINKLER: Okay, thank you. So recommendation for endorsement, how many on the committee vote yes?

How many no's? Any abstentions? All right.

Marlene, Ellen?

DR. MILLER: I'm a no.

DR. WINKLER: Okay.

DR. SCHWALENSTOCKER: I vote yes.

DR. WINKLER: Okay. Okay, it's 14 yeses, one no, one abstention. Okie dokie.

It's 11:15 or getting very close to 11:15, and we're expecting another measure developer to join us for measure 1448.

Is somebody from CAHMI on the line?

MS. REULAND: Yes, I am. This is Colleen Reuland.

DR. WINKLER: Hello, Colleen. How are you?

MS. REULAND: I'm good, thank you.

DR. WINKLER: All right. Measure 1448 is from work group two, and I'll give everybody a chance to pull that up, 1448.

This is developmental screening in the first three years of life. The description is, the percentage of children

screened for risk of developmental, behavioral and social delays, using a standardized screening tool in the first three years of life.

This is a measure of screening in the first three years of life that includes three age-specific indicators assessing whether children are screened by 12 months of age, by 24 months of age, and by 36 months of age.

And Marina, I believe this goes to you, yes?

CO-CHAIR WEISS: All right. Fortunately we have Colleen on the phone so that she can go ahead and make some comments about this particular measure.

And also, we've got with us Sarah from NCQA who can talk about it and its relationship to the next measure that we'll be looking at. So those two together seem to me to make sense to think of them as a pair.

1448, as Reva said, is a process

measure. It addresses the first three years of life. It's a developmental, behavioral, and social delays, standardized screening, and it's a process measure.

Beyond that, the rationale for the measure is that findings indicate that about 20 percent of children are screened in the first five years of life, but despite that evidence, the number that actually get into treatment who require it is relatively low. So this is intended to push behavior on the part of providers.

Let me go to denominator and numerator. Let's see, USPTF notes, I should say, that the evidence is not as strong as it could be. And so I would want to explore a little bit of that with both NCQA and also with Colleen.

And in particular, they note the absence of a focus on certain conditions such as autism. There is a rationale to that, but I'll leave it to the developers to talk about

that a little bit.

There are some -- there is a list of well-established and tested tools that can be used for this screening, which appears in your materials. And that list has been harmonized between the NCQA measure and the CAHMI measure as I understand it from the NCQA staff who are here.

Let me go to the denominator and numerator, hang on here a minute.

DR. SCHOLLE: Do you want us to comment on the autism issue?

CO-CHAIR WEISS: Yes, that would be great, while I'm looking for this.

DR. SCHOLLE: So, we actually have two separate measures. For NCQA we use developmental screening and autism, because they're different tools, screening tools recommended.

And so developmental screening can identify autism, but there are autism-specific tools for screening, so our panel recommended

splitting them out.

MS. REULAND: This is Colleen from the CAHMI. The recommended periodicity for developmental screening is different than for autism screening. And the measure -- we wanted this measure to map to the core measure that's recommended by CMS and SNAAC, and that core measure was anchored to the ABCD initiative, which was anchored to developmental screening.

And just to kind of note that the synergy -- the measure that we're talking about right now is a state-level measure, and then the measure that NCQA is going to be talking about is a physician-level measure, but they're in synergy for a certain age.

CO-CHAIR WEISS: Okay, so the denominator is broken down by age. Members who turn twelve months of age during the measurement year, members who turn twenty-four months of age, members who turn thirty-six months of age of the measurement year.

And the numerator -- I'll tell you in all candor, I had difficulty following this write-up. The other one was a little bit more clear to me.

Do you want to speak to that?

MS. REULAND: Sure. This measure is a hybrid measure, so for the numerator for claims data, it's a developmental testing or screening code, a 96110. And then for the medical chart, it's the documentation that date of screening that the screening tool was used and evidence that the screening tool was completed and scored.

We list criteria for the type of screening tools that meets the definition of developmental screening for risk for developmental, behavior, social delays, in terms of domains and in terms of specific attributes, in terms of reliability and validity. And then we list specific tools that currently meet that criteria.

The reason we approached it that

way is that we wanted to also allow for flexibility that if future tools were developed that met the reliability, validity, and developmental domain, that the measure wouldn't be constrained to tools that exist right now.

But we wanted to list the tools that meet the criteria right now at the same time.

CO-CHAIR WEISS: All right. So that's pretty much what I've got.

DR. ZIMA: I'd like to say that I think this measure carefully stays within the scope of the measure, and I thought it was quite wise to make the exclusion of autism, because I think that for our children who do meet the diagnostic criteria for autism or things like mild to moderate mental retardation, the implications for the types of services that they can access is very different.

And I think it would be more

complicated to interpret adherence to this measure if those more severely delayed children and more globally delayed children were included.

CO-CHAIR WEISS: So this is a care coordination measure, and it's also population health-based. And one of the things that probably should be pointed out is that there, of course, are disparity issues associated here, particularly in the income and financial arena.

DR. JENKINS: One question I had was in terms of the hybrid, whether the CPT codes in the claims data were reflective of the use of the validated instrument, or not.

And I completely understand the weakness of the codes there, but I'd just like to know the answer before we vote.

MS. REULAND: Speaking to the end part of that question, I want to make sure to address the question. I couldn't hear -- I know you said that you wanted to know about

the use of the claims code, but I couldn't hear the end of your question.

DR. JENKINS: That's my only question. That's my only question.

MS. REULAND: Yes. In terms of the claims, there's kind of state-level variability in terms of validity of claims, which is partially why we made sure that it was a hybrid measure to include claims and medical chart data.

And within the ABCD -- so this measure is building off the ABCD screening. Academy states experiences, and each of those states have kind of created state-level policies for what kind of codes could meet the 96110.

What codes or what kind of tools could be used in order for them to bill for 96110? When we had our stakeholder call with over 50 people from the screening academy states to review the measure, we did ask them if any of the states had done validity testing

of when a 96110 is billed, what do they see? And they had not done that work yet, and were planning -- some of them were planning to explore it.

I know through -- in Oregon, we're actually planning to explore the validity of it through our CHIPRA demonstration grant.

DR. JENKINS: But in general, most states, whatever their roles are, that allow billing under the 96110, only allow billing using an instrument that would meet the reliability criteria on the list of measures here?

MS. REULAND: The reliability criteria that we listed here was built off what a number of states in the ABCD screening academy had used in terms of a criteria that they laid out for what tools counted.

And because the SNAAC measure was supposed to build off the ABCD screening academy, we built off that work that they established.

So, for example, Minnesota, Illinois, Connecticut, I'm thinking of another state, all have this kind of criteria built into their state policies organ.

MS. BROWN: Can you all comment on the fact that the USPSTF sort of is not in favor of this? Is it that you feel that these other groups see it differently, or you're just more persuaded by them?

Or is the complex of services that that group reviewed, the Preventative Services Task Force, is that importantly different from what this measure is about?

MS. REULAND: I think what makes this measure complex, and also what makes it so valuable, is that it's identifying children at risk for developmental and behavioral social delays, and that some of those children will have developmental enhancement and promotion that actually pushes them to not have some of the developmental delays.

I think also one of the other

parts that is a bit complex is that the screening happens and then most of the -- a lot of the treatment happens in a different health system.

And so I think that's what's made some of the screening recommendations a bit complex, because you're talking about something that has a pretty long trajectory in terms of child health, and so therefore, your immediate outcome isn't as clear.

MS. BROWN: I understand that. I'm asking you to comment on the U.S. Preventative Services Task Force, who, at least in this brief, it says that they did not recommend this type of screening.

I'm just wondering why you disagree with them. And it could be because you have better data or different data, or because what you're recommending is different from the complex of services that that task force reviewed.

MS. REULAND: Yes. Sorry, We're

anchored to what is recommended -- the recommendations that are recommended by the Bright Futures, which cuts across all pediatric providers, and that is now part of legislation.

MS. BYRON: Just to be clear, the Task Force did not address it. So it's not that they reviewed it and rejected it. They just didn't address it.

DR. SCHOLLE: I'm sorry, they rely on Bright Futures?

No, those are separate. Those are separate. So Bright Futures is from HHS, right, you know Bright Futures.

So the challenge is that the Task Force evaluates issues where there's an evidence base to support it, and one of the challenges with child measures and preventive services is that there aren't randomized controlled trials.

There aren't -- and some of the outcomes that we're interested in, you'd have

to have like huge randomized control trials that look over 20 years to see what the outcomes are.

So that lack of evidence actually means that it's hard to do the kind of evidence review that the task force does.

Bright Futures looks at evidence -- I think it's called evidence-informed instead of evidence-based, because it does try to take into account what we know about the risk and what we know about the risk of developmental delays, that children are at risk.

If they're identified too late, they get into services late, and you've missed the opportunity for intervention, and so I think that's part of the challenge that we see here.

CO-CHAIR WEISS: To the extent that USPTF was involved in this at all, Sarah and Colleen, did they look at anything beyond speech and language?

MS. BYRON: So the U.S. Preventative Services Task Force did not address this. It doesn't mean they will never, and actually, a lot of different things go into what makes them take on a certain topic.

You know, they actually solicit feedback from the public to say, what are topics that you think are happening in primary care that you would like to request an evidence review for, and -- or they'll do kind of an environmental scan.

I don't recall where this falls on their prioritization list, but I think that we have to understand that they have not addressed it.

So they have not come out with a statement for or against it, and we really can't say anything about the evidence --

CO-CHAIR WEISS: So the statement, then, on page 5 is incorrect when it says that USPTF concludes and then goes on to describe

what the conclusion --

MS. BYRON: Let me -- yes, I'll have to look and see what that says.

I think, yes, they address autism. Let's see.

DR. SCHOLLE: The Task Force concluded evidence was insufficient to recommend for or against the use of brief formal screening instruments in primary care to detect speech and language delay in children.

However, this didn't address autism specifically.

Okay, so it's insufficient evidence, rather than not saying an A or a B or a D recommendation, insufficient evidence. So it's that issue about there not being enough evidence to make a decision.

MS. REULAND: Related to speech and language, but I don't think they reviewed developmental.

DR. SCHWALENSTOCKER: This is

Ellen. Can I just ask a question for clarification, that the measure listed as potentially a care coordination measure, and a couple of times people have mentioned early intervention, but I'm only seeing the measure as measuring if screening happened, not that referral or intervention actually happened. Is that correct?

CO-CHAIR WEISS: It's described, Ellen, it's described strictly as a screening measure. Is that correct, Colleen?

MS. REULAND: That's correct. Because it's kind of the first step to understand, were those kids screened who were identified at risk.

Then the next step would be care coordination, so it's the building block of it, but it doesn't address care coordination specifically, Ellen.

DR. SCHWALENSTOCKER: All right. Thanks, Colleen.

DR. WINKLER: Does everybody feel

you've discussed enough?

All right, for the committee, how many feel this measure 1448 meets the criteria for importance?

Any no's here? No.

Ellen, Marlene?

DR. MILLER: Yes.

DR. SCHWALENSTOCKER: Yes.

DR. WINKLER: So we'll move on to scientific acceptability for this measure. How many feel it completely meets the criteria?

Partially meets the criteria?

Minimally meets the criteria?

Marlene and Ellen?

DR. MILLER: Partially.

DR. SCHWALENSTOCKER: I agree, partially.

DR. WINKLER: Okay, thank you. Usability. How many feel it completely meets the criteria? No?

Partially meets the criteria?

Minimally meets the criteria?

Ellen and Marlene?

DR. MILLER: Partial.

DR. SCHWALENSTOCKER: Partial.

DR. WINKLER: Okay. And finally, feasibility, how many feel it meets the criteria completely?

Partially?

Okay. Minimally?

All right. And Ellen and Marlene?

DR. MILLER: Minimally.

DR. WINKLER: And Ellen?

DR. SCHWALENSTOCKER: Partially.

DR. WINKLER: Okay. Good. Now recommendation for endorsement? How many yeses?

Any no's?

How many abstains? No?

Ellen, Marlene?

DR. SCHWALENSTOCKER: Yes.

DR. WINKLER: All right, I heard yes from Ellen. Marlene?

DR. MILLER: Oh, I said yes.

DR. WINKLER: Oh, okay, I'm sorry. All right. 15 yes, 1 no, no abstentions. Okay.

So we need the next measure -- I think is 1399. This is developmental screening by two years of age. This is the measure from NCQA.

CO-CHAIR WEISS: Right. This is a process measure population health priority, and it relates not only to screening but also proper follow-up between six months and two years of age.

It's an NCQA measure, both publicly reported and intended to drive quality improvement.

It's described as fully developed and tested. You know the rationale behind it.

And let's see. We're looking at, obviously, there's a disparities issue. Here -- one study cited found that only 23 percent of low-income children receive recommended

preventive and developmental services. And of course, we know that there's been a chronic problem across the country with full compliance with EPSDT.

The USPTF did not review developmental screening generally. Rather, the Task Force reviewed the routine use of brief and formal screening instruments in primary care dealing with speech and language, and the recommendation received an I statement, which Sarah can expound upon if interested.

Let's talk about -- with the numerator and the denominator here, the numerator -- well, let's start with the denominator.

The denominator are children who turn two years of age between January 1st of the measurement year and the end of that calendar year and who had documented face-to-face visits between clinician and the child.

The numerator, children who had

documentation in the medical record of a screening for risk of developmental, behavioral, and social delays prior to the age of two.

CO-CHAIR WEISS: All right.

MS. BYRON: This is Sepheen. I just want to make a quick clarification, because I noticed a mistake in our form.

So, we updated 2A measure specifications to be documentation of screening and with the standardized tool. And as Colleen noted, we aligned this and harmonized the two measures together.

And that correction did not make it into the very beginning, where we actually said screening plus follow-up. So that part, in the very beginning of the form, is incorrect. But under 2A, it's correct.

CO-CHAIR WEISS: And I should underscore that there is an established and evidence-based list of tools to be used in this screening.

DR. JENKINS: So I guess the question is, is this measure exactly as the same as the middle measure that we just addressed?

CO-CHAIR WEISS: Not exactly, although there is some effort --

DR. JENKINS: Could I understand how it's different?

DR. SCHOLLE: It's complementary to the other -- so the other measure is intended for claims or chart review.

This is at a population level, state, or health plan level, and this is specified at the provider level for chart review reporting.

So it is -- this is the same as the middle part -- you know, one of the indicators in the other. But we're trying to specify it for a different reporting.

DR. PERSAUD: The denominator statement in this measure has that -- the denominator is children who have turned two

and had a face-to-face encounter, whereas the prior measure didn't specify that.

DR. SCHOLLE: Right. Well, it's for -- so that face-to-face visit is in the previous year, right, to establish that the provider has an ongoing relationship with this patient? So it's not requiring a visit during that year.

DR. CLARKE: But that's not when the screening takes place. Is that correct?

DR. SCHOLLE: The screening can take place any time between 12 months and two years, and any documentation in the record that the screening using a standardized tool was done, so that would allow if the -- even if -- I mean, a child's going to have a visit between 12 and 24 months, we'd expect, if they have that ongoing relationship.

But if they didn't, and somehow there was documentation from a daycare provider or somewhere else, that would count as well.

CO-CHAIR WEISS: And the prior measure went to 36 months, did it not, Sarah?

DR. SCHOLLE: The prior measure had three rates, one for in the first 12 months, one for between 12 and 24 months, and the other rate is between 24 and 36 months. So it's looking at screening according to the AAP recommendations.

This measure, we designed it to be part of that bigger composite measure that we're trying to build.

DR.CHEN: Is that why you picked two years instead of three years? Is there any evidence to support that two years is critical compared to three years of age?

DR. SCHOLLE: So, the reason for using that two years is because -- the reason we chose age two is because age two is when there's already measures that look at whether immunizations are up to date by age two. And so we're trying to fold in all the recommended services for children who turn age two. So

that's why we selected that age.

Yes?

MS. BERGREN: I just wondered if an exclusion should be a child that's already being -- already part of an early childhood intervention program.

MS. REULAND: It would be an exception.

DR. SCHOLLE: Okay, we did talk about that. So this becomes a challenge about how to document that.

In the record, and particularly in other -- in claims data where children may be in early intervention and the early intervention records aren't available to the claims, so if you were trying to do this from claims or from other records, you might not have that.

The way we would envision this being reported, rather than as an exclusion from the denominator, it would be recorded as an exception by the provider that is reporting

it. Since it's a provider-level measure, it would be a way for the provider to say, this child doesn't -- this measure doesn't apply because the child is already in services.

DR. JENKINS: I'm sorry, what's the difference between an exception versus an exclusion?

DR. SCHOLLE: So, actually, part of this -- an exclusion is this service is never appropriate for this child, okay? And you remove them from the denominator.

So one of the discussions that we had was, well, you know, there might be some children who are in services for whom doing this screening and having some sort of updated, are they on track, where are they, makes sense.

It's not exactly primary screening anymore, but the -- our panel felt like it made more sense to -- not to exclude children, but rather to offer that exception.

So an exception is that for this

particular child, it doesn't apply. Exclusion, it never applies to that child.

MS. BROWN: So the relationship of this measure to the previous one, as the previous one said, developmental screening, three times. And this one says, on the second one, the second year of life, if something is found, there needs to be follow-up.

Oh, that's a mis -- okay, so what's the difference between part B of the previous one and this one?

DR. SCHOLLE: The level of specification -- the previous measure is specified at the population level, for health plans or states, and this one is specified at the provider level.

DR. JENKINS: And the last one was also had the option to look at claims data, and this one is exclusively based on chart review.

MS. BROWN: So what does that mean? I don't --

DR. SCHOLLE: I think if we just think about this one, what we're thinking of it as is as an accountability measure at the practice level for all children who meet the denominator criteria, which is essentially based on a face-to-face visit in a certain time window and having a birthday in a certain time period, and having the screening by the instrument documented in the record as having been performed.

MS. BROWN: So, what's the rationale then for only requiring a provider to do -- to take action in the second year? What about the first year and the third year? Is it that we're just choosing one because three requirements is too much? Or I don't --

DR. SCHOLLE: So this goes back to the logic of our -- and I apologize, you know, the logic of the composite measure that we were developing, and it seems like we really confused things by presenting the measures

individually in addition to presenting the composite.

But the logic of the composites, that Sepheen presented yesterday, and I think that you'll review later today, is that we wanted measures that would look at the comprehensiveness of preventive services for children at key developmental ages, six months, two years, six years, thirteen, and eighteen.

And so all the measures you're seeing that are not currently HEDIS measures, we're presenting them for those age windows.

And so we chose age two because there's already measures that look at children by age two, and we wanted to try to fold that in as a critical time frame for doing it.

So there were some concerns on our panel that age one is a little early to be doing the screening, and that age two -- by 18 months is really the critical time to do the screening, and so that's how our panel said,

focus on age two.

DR.CHEN: Yes, and I agree with that. I respect that as well. I'm a little bit worried as an accountability measure, age two is a little bit too critical.

I mean, I think most people screen these kids at eighteen months, probably. That would give you basically six months to get that in as an accountability measure.

I would be much more comfortable, I don't know about other people at the table, with age three. But I know that doesn't fit into your framework of the composite measure, which I understand.

But I'm just a little bit uncomfortable at age two. I think it's too high a bar, especially for immigrant children and foreign, bilingual children. You can't pick up speech delay in these kids at age eighteen months. I think it's just too high a bar.

MS. CARLSON: This is maybe more

of a technical question as well, but it looks like you're only including physician services, physician screens. What about mid-levels, advanced practice nurses?

DR. SCHOLLE: We did not intend to exclude mid-levels.

MS. CARLSON: So that's just an oversight on the -- okay, thank you.

DR. ZIMA: Just to clarify, this is for any type of health plan. It's not restricted to public --

DR. SCHOLLE: I said, this isn't a health plan measure, but no, it's not restricted by insurance. It's intended for all children.

CO-CHAIR MCINERNY: Okay. Good discussion.

DR. WINKLER: Anything else? Ready to -- so for measure 1399, developmental screening by two years of age, this is a provider-level measure, not a population and health plan measure.

Of the committee, how many feel that it meets the importance criteria?

DR.CHEN: We're voting at 24 months, right?

DR. WINKLER: Unchanged.

DR. WINKLER: How many no's? Zero.

Abstain? One.

Marlene and Ellen?

DR. MILLER: Yes.

DR. WINKLER: Ellen?

DR. SCHWALENSTOCKER: Yes.

DR. WINKLER: All right. So under scientific acceptability, how many feel it completely meets the criteria?

Partially meets the criteria?

Minimally meets the criteria?

Marlene, Ellen?

DR. MILLER: Partial.

DR. SCHWALENSTOCKER: I agree, partial.

DR. WINKLER: Okay. So,

usability, criteria, completely meets?

Partially meets?

Minimally?

Marlene and Ellen?

DR. MILLER: Minimal.

DR. WINKLER: Okay, Ellen?

DR. SCHWALENSTOCKER: Partial.

DR. WINKLER: Okay. And on the feasibility criteria, completely meets?

Partially meets?

Minimally meets?

And Marlene and Ellen?

DR. MILLER: Partial.

DR. SCHWALENSTOCKER: I agree, partial.

DR. WINKLER: Okay. So, recommendation for endorsement, all those for yes?

All those for no?

Abstain? Two.

Marlene and Ellen?

DR. MILLER: Yes.

DR. WINKLER: Ellen?

DR. SCHWALENSTOCKER: Yes.

DR. WINKLER: Okay. 13 yeses, no no's, and two abstentions. Okay. Wow.

So we might as well finish this last one in the group, measure 1341, autism screening. This is, again, from work group two.

DR. MILLER: This is me.

DR. WINKLER: Right, this is Marlene's. And the measure is the percentage of children who turn two years old during the measurement year who had an autism screening and proper follow-up performed between six months and two years of age.

All right, Marlene?

DR. MILLER: So I'll just sort of talk through my sense of looking through the measure evaluation.

As you've said, that is the measure. It actually sort of implies two components. It's the percent of kids who had

a screening, which is very well defined in the measure specs, and then proper follow-up performed, and that is actually very poorly specified in the definitional aspects of this measure later on through it.

This measure has been tested. There is data submitted on about 180 charts, but it has not actually been used in practical purposes, and it is being put forth for public reporting, internal QI, and also accountability.

I think we all know that sort of evidence behind it, that autism is an important problem, that it's prevalent, and that perhaps we do not identify these children as early as possible.

That being said, there is a key paragraph on outcome or evidence to support the focus of the measure, which is sort of the Achilles' heel on autism is it's sort of, what is the appropriate intervention, particularly at younger ages, is unclear, and the measure

review articulates that.

So we may want to screen, but we don't actually know what to do once we know that data, which is a problematic, to me.

I think that's part of the reason on the summary review that the U.S. Preventative Services Task Force did not recommend this. They said not for or against a brief formal screening tool to detect these types of delays in it.

I think a larger issue, problematically, I have, is that there -- there is not actually one tool recommended. There is a list of tools, and there's no data provided on the varying list of tools that you could use to screen on what the various sensitivity and specificity is of the tools.

So, it's not -- you can use any one of these, but we don't know if one is better than the other, and it sort of gives you credit for using any and all of them.

The review of process, if you

will, is a chart review, which also brings up the burden. There's a note at the end that NCQA is considering looking how, eventually, how to specify this measure for electronic health records, but this right now is the burden of the chart review for it.

There is no risk adjustment, which is probably appropriate. However, there's also no disparities built in, which I'm not sure is appropriate, that there might well be some disparities of care, and it could be easily built into the measure, but has not been done.

There -- also, there's been no reliability testing. Remember I told you on the testing results, there's about 180 charts that were looked at, but there was no reliability testing, because it's a very subjective of looking through the chart, was one of these tools documented. And then, remember that second part of the measure is, was proper follow-up, that has a pretty broad

aspect of what they consider proper follow-up.

To read the sentence, it says, for abnormal or indeterminate results, evidence of a confirmatory test, referral, or treatment, and is no better specified than that of what is a quote unquote appropriate follow-up.

Validity testing was only expert panel, sort of on face validity. And again, there's no mention of the link to outcome, meaning, that if I do this screen, I know that children do better.

And I think that hinges again, going back to U.S. Task Force and the fact that there's not necessarily a proven intervention that's known for this. I think that would probably be all of my summary of looking through this measure.

MS. BYRON: This is Sepheen. Could I make another clarification? So, this measure, actually, we structured similarly to the developmental screening measure, and it should be structured the same.

We had an issue of different deadlines for each of these, and we actually -- NQF was kind enough to give us an extension to work with CAHMI to make sure that we had harmonized our developmental screening measure, but I believe we turned in autism first before that.

So, the form is incorrect. It really shouldn't be with follow-up. It should be screening with a standardized tool in the medical record.

DR. SCHOLLE: And the other clarification is, in terms of the tools, we use the list of tools that are recommended, I think, by Bright Futures.

We did not require a specific -- we didn't say, use this specific tool, because there are multiple tools out, available, so we relied on those existing guidelines for which tools are available that were recommended by the AAP.

DR.CHEN: So, I just want to make

a comment. So, if I had a problem with the previous one at two years of age, I have a lot more problems with this one at two years of age.

The Preventative Services Task Force did not find any evidence or insufficient evidence to support even up to five years of age.

And for autism spectrum disorder, it's a very heterogeneous disorder, both genetically and environmentally, as well as behaviorally.

I can't expect anybody to make any type of assessment, even in a validated screening tool, for age two. And that's just my opinion on that.

DR. MILLER: And just so I can clarify, the measure is that you do the screening anywhere between six months and two years of age.

DR. SCHOLLE: Right, and the recommendation from Bright Futures is at the

18 month and 24 month only. It doesn't go up to age five or anything, it's just at 18 and 24 months that it's recommended by Bright Futures.

DR. MILLER: I know, but that's not what this measure specifies.

DR. SCHOLLE: This measure would count anything that happens before the -- the reality is, I don't think anybody would do an autism screening at six months, but we count whatever -- remember, this is part of a composite, and in that composite, we're looking for events that happen over that eighteen month period.

So if the panel would like for us to limit it, then we would. We're trying to anchor to that second birthday, and the Bright Futures recommendation says 18 months and 24 months, so usually we give a couple of months to allow for something to happen.

So, if they're recommending it at 18 months, then, you know, if you give people

wiggle room for patients not making appointments by 24 months would be at -- I recognize the issues around language delay, that Dr. Chen recommended.

DR.CHEN: I'm sorry, this question is for Marlene.

Do we know if there is any evidence in terms of specificity and sensitivity of the evaluative screening tools at age two for autism? I mean, just how accurate is that screening tool?

DR. MILLER: I don't know. You're going to go beyond my expertise. I will say there's one, two, three, four tools that are listed. And then there is a sentence that says, because of lack and sensitivity and specificity, the Denver and the Revised Denver are not recommended.

So there's a suggestion that there is, someplace out there, this truth is known, but there's a long list of tools that are recommended, but not information about their

sensitivity and specificity.

DR.CHEN: I'm primarily worried about labeling. I mean, if it's not -- if it's sensitive but not specific, which most screening tools should have as a good screening tool in terms of an instrument, then you're basically labeling these kids at two years of age as having ASD if you screen them. I mean, if it's a false positive.

(Off-mic comments.)

I know, but --

MS. BERGREN: But then you refer them to the school system, and they take over. And they don't label them that young.

UNIDENTIFIED SPEAKER: What's the false positive rate?

DR.CHEN: Let me just give you an example. My own son, right? I made a mistake of filling out his elementary school application that we are bilingual at home. He now has a mark as an English not proficient student, until he tests out of this label.

He's never been evaluated as being English non-proficient. It's the state of California's requirement because there are so many immigrant children that we have to address kids that are English not proficient.

So, we speak English at home, but we are bilingual in terms of heritage and language proficiency. But my son has this mark of English non proficient at elementary school, because I filled out the form, instead of having English as our language at home, I filled out bilingual.

So, I mean, this is just language. It's not a big deal. But I see a lot of kids with autism that -- in the charts, that they're screened, positive for ASD at two years of age. And that worries me.

DR. ZIMA: I'd like to provide a little context, too, and that is that I hear some discussion combining ASD with autism.

And there are two issues. One, the diagnostic criteria for autism, if you

noticed, is really still quite broad and underdeveloped, and even children who meet that diagnostic criteria for autism can have wide variation in the level of functioning and whether they're also retarded.

And then if we broaden it a little bit more to ASD, we get an even larger, exponentially greater heterogeneous group of children. And so I think in some ways, what my concern is that this is a little premature, given the limitations of how we make these diagnoses, versus, like the other indicator, where I think there's a lot more established evidence as far as screening for a global developmental delay.

DR. MILLER: And I will say some of the background evidence in the measure work group sort of bounces back and forth between autism and ASD, so CDC recommends it for ASD.

Some of the other language is for autism screening tools. I'm not sure it's well specified exactly. The measure is named

autism screening, but a lot of the language behind why and where guidance, if you will, is pulled from, is about ASD.

DR. GLAUBER: And certain states, I don't know the exact number, but most recently, Massachusetts, have legislatively mandated insurance coverage of autism treatment services, so that creates its own provider community and momentum towards early diagnosis and treatment, and the potential that kids may get falsely labeled early and get on a treatment path that may not be necessary.

MS. BERGREN: I hear the concerns about the labeling, but having worked in that realm of the under six for the school districts, it's such a difficulty to get the kids in the system, and then we monitor them until they either meet the criteria or require services or don't. And we work with the primary care providers on that.

I guess my concern is, the delay

that would occur in getting them into the early intervention system, that this mechanism allows that early referral and then constant monitoring. The onus is on the school to keep tabs on that family and myself, we did a lot of home visits to do that. The onus wasn't on the parents to get themselves to us.

CO-CHAIR MCINERNY: Okay.

DR. WINKLER: Any further discussion? Ready to decide?

DR. JENKINS: I'm still confused. Could I just ask a measure developer one more time without referring to Bright Futures or someone else, just the measure developer themselves, why they believe that screening with these sets of tools at two years is an important performance measure at the accountability level, given the current state of knowledge for primary care providers?

DR. SCHOLLE: So, I'll try that. So we worked with a multi-stakeholder process to try to identify measure concepts that are

important. And what I can say is that in our process, that included physicians and Medicaid directors and consumer advocates and health plan representatives, they prioritized this measure.

We reviewed the evidence base. We provided them information about the guidelines, the various guidelines, the recommendations, the expectations from different states. And the panel recommended that we include this as a separate measure from the developmental screening measure.

So, it was their sense of the importance of this problem, the need to identify children early, and an opportunity to do that in the primary care practice, and to include that as an important component of well care for children.

DR. JENKINS: So, just to clarify, does that mean that they believed that children should be identified as potentially having this problem at age two years, and that

there were valid instruments to accomplish that that could be applied in a primary care setting?

DR. SCHOLLE: Yes. I think they were looking to the recommendations from the American Academy of Pediatrics that say, screen children for the risk of these disorders, using these tools.

DR. QUIRK: How do the tools perform?

DR. SCHOLLE: So, I'm sorry, I don't have the information about the sensitivity and specificity of those specific tools.

DR. QUIRK: Yes, because sensitivity and specificity would vary according to how frequently you'd expect to see it in a population, and positive predictive values would be more useful, number one.

Number two, we don't have a validated tool --

DR. SCHOLLE: Okay, I take it back. I -- we did. We do have that information, so the CHAT, the Checklist for Autism in Toddlers, I believe that's one that's commonly used as a sensitivity that's low, 0.38 to 0.65, and a specificity that's high, 0.98 to 1.0.

And it looks like there are other measures. The M-CHAT has moderate sensitivity and high specificity. There are a couple of other tools.

So you can see there, when we reviewed this, we did have these tools available to our panel. So I apologize, it's not an area of expertise for myself, but it appears that there are tools that accomplish that.

DR. MILLER: So all tools are not created equal. Why would the work, the development work of this, not identify the best tool and then put that forward?

DR. SCHOLLE: Well, that's a

challenge that we have with all of the measures that we've talked about.

And in earlier discussion, in some cases, we said, hey, if we just get them to talk about it, we're thrilled. And it really depends on the extent to which a measure -- specific tools have distinguished themselves as the best tool, and where there's consensus that there is a best tool.

And it also has to do with the cost and availability of the tool. Some of the best tools actually have prices attached to them for use, and I can't remember which measure that was.

So our panel took the approach that for the measures where we're requiring a standardized tool, that we know that different states have different requirements about which tools they'll pay for, and which tools they want people to use. And different providers have different preferences for tools.

So what we did is we worked with

our panel and used the recommendations from the academy to say, these are tools that have some evidence of -- that they work well, and we said, use one of these standardized tools.

It's been our experience that if we say, use a specific tool, that people that have other tools, are using other tools, aren't very happy.

And so it's a challenge, because we're trying to move towards getting people to do something in a standardized fashion. And if there's real consensus that there's only one tool and it's publicly available and everybody can use it for free, then that would be best, but that's not always the case.

DR. JENKINS: But you do understand that even if there's a short list of tools, that you're actually advocating strongly by being here for universal screening and application of these tools to all two-year-olds in the United States.

DR. SCHOLLE: Yes.

DR. JENKINS: That's what you're actually asking us to endorse.

DR. MILLER: Right, although the data is not all there in terms of showing that if you do do this screening, that there is absolutely improved outcomes for children with autism and that there's absolute clear interventions.

DR. GLAUBER: And I think this is an area where, unless you're aware of it, there is the potential for harm.

And we'd like to see some evidence that kids who screen positive, who may, upon further evaluation, either fall into an indeterminate group who have heightened monitoring, whether the implications for that child and that family's functioning, I think that would be a concern about endorsing universal screening at this age.

DR. QUIRK: I have a problem, because without really knowing what the performance characteristics of -- for any

tool, for the variation amongst tools, if you're going to allow a choice of a menu of tools, that there is an enormous social burden, and an enormous financial burden, at perhaps a time in development where the sensitivity, specificity, the performance of the tool would be better applied later, perhaps. But that's a health services research project.

DR. WINKLER: All right. So, for this measure, 1341, autism screening, how many feel that it meets the importance criteria?

And Marlene and Ellen?

DR. MILLER: Yes.

DR. SCHWALENSTOCKER: Yes.

DR. WINKLER: Okay. How many say no? Four, okay.

Okay, all right. In terms of the criteria for scientific acceptability, how many feel it meets the criteria completely? Zero.

Partially? Two.

Minimally? Five.

None at all?

And Marlene and Ellen?

DR. MILLER: None.

DR. SCHWALENSTOCKER: Minimally.

DR. WINKLER: Okay. And usability, completely meets?

Partially meets?

Minimally meets?

Not at all?

And Ellen and Marlene?

DR. MILLER: Minimal.

DR. SCHWALENSTOCKER: I agree.

DR. WINKLER: I can't hear you, Ellen.

DR. SCHWALENSTOCKER: Minimal.

DR. WINKLER: Okay. All righty.

And now for feasibility, completely meets?

Partially meets?

Minimally meets?

Not at all?

Well, okay.

And Marlene and Ellen?

DR. MILLER: Minimal.

DR. WINKLER: Ellen? Ellen, are you there?

DR. SCHWALENSTOCKER: Yes, I said minimal.

DR. WINKLER: I'm sorry. Can't always hear you.

Recommendation for endorsement, how many yes? How many no?

And Marlene and Ellen?

DR. MILLER: No.

DR. WINKLER: Ellen?

DR. SCHWALENSTOCKER: I guess I have to say no.

DR. WINKLER: Okay. All righty. Were there any abstentions? Okay.

All righty. So that's that. We probably could do public comment, and then break for lunch.

Operator, would you see if anybody

might be listening out there who would want to ask a question or make a comment?

OPERATOR: Certainly.

If you have a question or comment, please press star one at this time. Again, that is star one for a question or comment. We'll pause for just a moment.

DR. WINKLER: Thanks. Probably nobody's out there.

OPERATOR: There are no questions or comments at this time.

DR. WINKLER: Thank you. Anybody here in the room, beside the committee, want to say anything at this point, or is everybody ready for lunch?

Okay. So we will reconvene at maybe 12:40, since we're just a little bit early. Sound good?

Thanks.

(Whereupon, the above-entitled matter went off the record at 12:10 p.m. and resumed at 12:45 p.m.)

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

12:45 p.m.

DR. WINKLER: We skipped over one measure from this morning's agenda, and that's 1404, lead screening. And why don't we go ahead and do that one, and then go through the rest of the afternoon's agenda.

So, 1404, and Shannon isn't here, so let me go back and grab it.

I can tell you this is not the first time NQF has seen this measure. This measure came through and had an interesting history, so, this is a measure of the percentage of children two years of age who had one or more Venus blood tests for lead poisoning by their second birthday, and the -- - can you scroll down to the specifications just so everybody sees them?

The numerator is at least one capillary or Venus blood test on or before the

child's second birthday, for all children who turn two years old during the measurement year, though I believe, and this is where it gets confusing, this is just for Medicaid children, correct, guys? Yes.

So this measure is a straightforward measure. It's been around for a while. It's a HEDIS measure at the Medicaid health plan level, correct?

This measure was reviewed in a project we did about a year and a little bit ago, and it was originally recommended.

However, it was out for public comment during August of 2009 when the CDC released changes in the recommendation for screening for lead. And those recommendations basically advocated, rather than assuming, like, for instance, a Medicaid population would be automatically high-risk, would be to look at local conditions and local risk factors rather than something that's more blanket like this.

And as a result of that change in recommendation from the CDC, the steering committee changed their recommendation to recommend that this measure go forward for endorsement. So, like I say, it was because things were just rapidly changing right at that particular time.

So the other folks in work group whatever, three, who looked at this measure, I'd be interested to hear your thoughts in terms of --

DR. RAO: I think there's a couple of considerations. One is the rapidly changing recommendations from the CDC. The other thing is the overall significance of modestly elevated lead levels.

You know, I live in a city that's got lots of lead and lots of older housing and stuff, and it's very rare that we see children with elevated lead levels that require some sort of intervention, and so I just think we could probably spend resources more wisely.

CO-CHAIR MCINERNY: Is this for all children, or just those on Medicaid? Is that what the measure says?

DR. WINKLER: Yes.

CO-CHAIR MCINERNY: I'm having trouble pulling it up, but, Medicaid only. Okay.

CO-CHAIR WEISS: I assume that means Medicaid and CHIP, right? Or does it?

CO-CHAIR MCINERNY: Well, it varies from state to state. In some states, CHIP is a Medicaid expansion, and in other states --

CO-CHAIR WEISS: Right. It's about a third, a third, a third, I think. In about a third of the states, it's indistinguishable from Medicaid.

About a third where it's a mix, about a third where it's a separate program. So would this screening be for the CHIP kids as well?

CO-CHAIR MCINERNY: It's probably

a state to state.

In New York, it would not be. CHIP is separate from Medicaid in New York.

DR. JENKINS: I'm not seeing a Medicaid restriction in the numerator and the denominator statement.

CO-CHAIR MCINERNY: Okay, I missed that. Thank you. Thank you.

MS. CARLSON: The problem that we have in Wisconsin is that our state Medicaid Department of Health Services program doesn't really accept or recognize the CDC recommendations for blood lead.

And in fact, they've made this a priority for the state, and they don't even accept this NCQA measure. They tightened it up so they want to see a blood lead at one year of age and a blood lead at two years of age.

And they don't accept the verbal assessment, which used to be, I think, the guideline in place. You do the verbal

assessment and then test, if you need to, after that.

This is one of the hardest measures for health plans to perform on because there is so much disagreement in the medical community and not a lot of commitment in the medical community to this testing schedule, so.

DR. GLAUBER: Yes. This is, in my view, an example of where the health plan has accountability with very little ability to impact improvement if, indeed, improvement is needed, because of the lack of support among providers for doing this.

MS. CARLSON: The other issue is the data isn't always easily accessible. For instance, in Wisconsin, WIC programs can perform blood lead testing, health departments can perform blood lead testing. Then the state has a state database because it's legally required that the results be submitted to that database. But it's not 100 percent

complete, and they don't match their Medicaid population at 100 percent with that database, so it is problematic to try to meet this metric no matter how they test for it.

DR. JENKINS: It looks to me like it's also being offered at the clinician level. Could I ask a measurement developer to speak to some of these issues?

MS. BYRON: Yes. So, this is actually as acknowledged, it's a long standing HEDIS measure at the health plan level.

We also considered it in our comprehensive well care composite framework as by age two, and a lot of these issues, they've come up in our measurement advisory panel.

I mean, I have to say, this is one where some people are really for it, some people are really against it. In the end, we did leave it in so that we could field test it and so it is at the physician level as well.

DR. GLAUBER: Also, in my experience, even when you have physician

support and ordering of the test, there's a fair degree of parental non-adherence to actually getting the test done, because it's a blood stick and it's often in conjunction with visits in which the child has gotten several immunizations.

CO-CHAIR MCINERNY: Well, to address that, in New York state -- and unfortunately New York state still says, for all children, not just Medicaid children, but all children in New York state.

And what we've done in our practice is we actually have a machine that does the lead test, and you do a finger stick in the office, get the result, and that way you know it's been done. And that's worked reasonably well.

I personally think it's not a very good use of resources, but to keep us from violating a state law, we go ahead and do it.

And I think the problem is that the real push should be for primary lead

abatement, and fortunately, Rochester has been very strong in that area.

And we now have narrowed it down to barely 1,000 children in the whole city who have mildly elevated lead levels. And finding someone with a lead level over 15 is pretty rare these days, even Medicaid children.

If there's no further discussion, we'll take a -- oh, Alex?

DR.CHEN: Quick question. Do we know what AAP's stand on this is?

CO-CHAIR MCINERNY: I think it's high-risk.

DR. LIEBERTHAL: I'm just trying to picture the periodicity schedule in my head, and I think it's one of those that has a star, which means based on risk assessment.

MS. BYRON: I have it in front of me, and you're right. It has a star, so it's a risk assessment.

CO-CHAIR MCINERNY: Right. So it's not the test itself, it's just, ask a few

questions.

DR. QUIRK: It says that -- it's on -- where am I, 1-B, where it's got AAP 2005, it says, all Medicaid eligible children must be screened, and then it tells you what to do with the result of the screening. That's 2005.

MS. BYRON: Yes, and maybe -- we found that those recommendations had just come out at around the time when we were developing this and refining it.

And this puts all of us, I think, in a difficult position. And the feedback that we receive from organizations like the CDC, I don't want to speak for the CDC, but what we did hear is, this measure, at least, because there are some requirements for it, offers a standardized way to do it because it's specified.

You know, we find that some states require it, others do not, and so we're all kind of caught in between everything that's

going on with that.

DR. LIEBERTHAL: If some states require it and some states don't, it's up to the state, then, to decide if they're going to monitor or do a performance measure on compliance with the state law.

But you can't generalize from what individual states are requiring if it's no longer a supportable, evidence-supported measure.

DR. MILLER: This is Marlene. I think I sort of echo that. I think this highlights with all the struggles and the decreasing incidence, if you will, is it, not everything that can be measured should be measured.

DR. GLAUBER: Just a question to NCQA, have you seen any movement in this measure in the time that you've been tracking it?

MS. BYRON: In 2008, the rate for Medicaid plans was 66.7. In 2009, it was

66.4, so.

DR. JENKINS: But do you know the numbers of children with especially high lead levels that were identified through the screening?

MS. BYRON: We don't track that information. I think the CDC has that information, though.

CO-CHAIR MCINERNY: Okay. If there's no further discussion, we'll take a vote.

DR. WINKLER: All right. How many on the committee feel this meets the importance criteria?

Yes?

No?

Any abstentions? Is that what I'm seeing, one?

How about Marlene and Ellen?

DR. MILLER: No.

DR. SCHWALENSTOCKER: Agree, no.

DR. WINKLER: Okay. All right.

Okay. We're continuing on our afternoon agenda, so the next measure to discuss is from work group two. It's measure 1397, Sudden Infant Death Syndrome Counseling, another measure from NCQA.

Dr. Miller, you are --

DR. MILLER: It's me.

DR. WINKLER: Right, it's you.

DR. MILLER: Should I start?

DR. WINKLER: Sure, go ahead.

DR. MILLER: Okay. So this is a comparable, in a way, process measure to that autism one we discussed where the goal is to report on the percent of children who turn six months during that measurement year who had a Sudden Infant Death Syndrome counseling and proper follow-up.

Some of the comments I made earlier about defining proper follow-up also happen here. It's not very well specified in the measure, although it is in the name of the measure.

I think in terms of importance, we all know Sudden Infant Death has affected large numbers of people, and the measure evaluation nicely goes through all the evidence.

So, for example, based on all the back to sleep campaign in the `90s, the rate of infants that are placed in the prone sleeping position decreased by 64 percent in a recent survey.

So 75 percent of kids used to be and with all the back to sleep campaign, that's down to 11.3. That's looking at data through about 2002, so it's still relatively fresh.

So there's lots of guidelines about the fact that people should be counseled on this, and I think that's unquestionable.

What is funny about this measure is that there actually is no tool recommended, and so I think that's because the guidelines have it.

But there's not a checklist created yet of what should you counsel on, what should be the things you check about. And in fact, I think partly because there's not an actual tool sort of evaluated out there, this has never been, as far as I can see, from the evaluations brought forth between the U.S. Preventative Services Task Force, because its guidelines and my gathering is vaguer language about, this is something you should counsel on, but there's not set screening items for it.

So the lack of the tool, quite honestly, is my biggest problem with this. There is nothing. The numerator would detail a subjective chart review where you would document, engaged in discussion about placing infants on their back, and they would check if that was documented.

It talks about a checklist indicating that SIDS was addressed, although there is not any tool provided with this, that

there is counseling or referral for SIDS education, that the member receive educational materials on SIDS, although again, there's not set tools out there. Find that anticipatory guidance for SIDS was given.

As I mentioned, it's all chart review subject to finding those elements, and so there is a high amount of burden.

There has not been any reliability testing on this. Again, with a -- very comparable to the autism one, about 190 records have been looked at, and there's no reliability testing on two people look at the same chart and can they glean the same information on it.

There is no risk adjustment in this measure. In their sample for meaningful difference, again, here, they say they looked at 180 charts and almost 80 percent of kids had some documentation of it, which bears out the question of, is there enough of a signal there?

There is no disparity stratification, and again I would question that there needs to be based on certain lines in terms of risk for SIDS, but that's not in here.

Let me see if there's anything else that I would add on the measure. I don't think so. Hang on. I think that covers most of it.

CO-CHAIR MCINERNY: Well, just another issue, in the title it says, and follow-up, proper follow-up, but the measure -- this numerator does not say anything about follow-up.

DR. MILLER: No. Proper follow-up is a counseling or referral or educational materials, but they're not further specified of what that is.

MS. BYRON: This is NCQA. So this is a -- this is just a straight counseling measure. We wanted to see that counseling was documented in the medical record, and what

you're seeing in terms of counseling or referral, those are -- again, this is structured similarly to any other counseling measures for existing HEDIS measures.

So it's a list of anything that would count as a numerator hit towards the measure.

So we say that if you see any of these five things, engagement about, discussion about placing infants on their backs, a checklist indicating that SIDS was addressed, because often times, physicians will have a checklist that just says, did you ask about, A, B, C, and D, and SIDS would be one of them.

If there's any counseling or referral for SIDS education, that would count as a numerator hit. It would mean that the physician counseled the patient.

If there were educational materials received or there is documentation of anticipatory guidance, so that's the

guidelines that we give for meeting that numerator.

DR. SCHOLLE: Again, we apologize for the inconsistencies. Part of this is that the title of the measure was pulled from what we tested, and -- because we intended all of these measures to be counseling and follow-up or assessment and follow-up.

And what we found was that was very hard to document in the chart reviews, and so that's why in our -- in what we presented to you, it's what we think is feasible for chart review.

DR. JENKINS: Can I just ask a question about the timing? And I understand that the six-month time frame was chosen by the NCQA task force, but in terms of the proper time for counseling, it seemed a little late to me.

It seemed similar to some of our neonatal discussions we had yesterday.

DR. RAO: I just want to agree

with Kathy. I think it should take place around the time you discuss breast-feeding, probably before the baby is born, not after they've been sleeping for a few weeks.

DR. QUIRK: I just have a question. Wouldn't, in term of the timing of this counseling, not to throw it back to the hospital, but isn't this something that ought to occur before the baby leaves the hospital?

You know, there's all those things, and that should be documented on a discharge note in the hospital by the nursery pediatricians or the nurse practitioner. That's when breast-feeding is getting done.

DR. GLAUBER: And if the intent of the measure is that counseling occur to prevent the -- reduce the likelihood of SIDS, I think a key component of it should also be assessment of ETS exposure in the home, since that's a major risk factor, and that's another measure that we're going to consider. But to be a really inclusive measure, I think this

ought to be a component of it.

CO-CHAIR MCINERNY: Any further discussion? We vote?

DR. WINKLER: All right. So how many on the committee feel this measure meets the importance criteria? Yes?

Any no's?

Abstains? Two.

How about Marlene and Ellen?

DR. MILLER: Yes.

DR. SCHWALENSTOCKER: Yes.

DR. WINKLER: Okay. All right. So, for scientific acceptability, how many feel that it completely meets the criteria?

DR.CHEN: I'm sorry, can I just make one comment about scientific acceptability?

So, there is evidence that sleep position is associated with SIDS. But that's really where -- and smoking, of course.

But there's -- that's not 100 percent of all SIDS. That's only some

percentage of SIDS. So there's evidence there, but there's no evidence anywhere else about the rest of the 20 percent of SIDS that we can't correct by changing sleeping position.

So I think it's hard to vote on the scientific acceptability, because SIDS is a garbage wastebasket thing, where there's multiple diagnoses within that, we just couldn't find out why they died.

But a lot of it is sleep position, and if that's what we're voting for, there is plenty of evidence. But if there is all SIDS, then I'm not sure.

DR. MILLER: Well, I think what we'd be voting on is not so much, does sleep position matter, but does a counseling episode in your pediatrician's office influence your likelihood of not doing prone sleeping?

DR.CHEN: Well, counseling helps with sleep position, but not with anything else. But I do agree, six months is not very

--

DR. SCHOLLE: If I can clarify, because if the issue is that by age six months is too late, age six months is the sampling approach, children who turn six months.

We can set the measure as, was this discussed, if you think it's appropriate for it to be discussed by the pediatrician or the pediatric provider by -- at the first pediatric visit, then we could frame it that way, okay, because all the children have to have been with this provider since birth.

So if that makes more sense, then we could set that time frame and say, it should be done at the first pediatric visit, if that's what the committee would like to see, if it's really about timing.

I also heard some people say, you know, that's too late. It should be done during pregnancy or at the hospital.

Our panel actually reviewed this measure. They reviewed depression screening,

they reviewed breast feeding counseling, and on breast feeding they said, yes, you know, there's nothing for pediatricians to do. We field tested it and it came back and everybody's already into whatever they're going to do by the time the pediatrician or the pediatric provider can have any influence.

So, but they saw this as being something where the pediatric provider had a responsibility to reiterate this advice, so I wanted to see if the panel felt we should go back and really re-think the timing.

DR. JENKINS: For me, the answer would be yes.

DR. GLAUBER: For me, your important word was reiterate. Yes.

DR. MILLER: I would say yes, and I would want to see all new validity and reliability submitted with that, because that's a different measure.

DR. SCHOLLE: We actually tracked the timing of the counseling in our field

test, so we may be able to provide some information related to that.

DR.CHEN: But I think that's for good validity data, just by the fact that SIDS dropped by significantly right after the back to sleep campaign was initiated.

I mean, I don't think you need any more validity data than that, if you are only addressing SIDS that's addressable by changing sleeping position.

DR. CLARKE: I think the absolute deadline for the first knowledge of the sleep position by the parents is hospital discharge. It's got to be before that.

DR. WINKLER: All right. So you did feel that it meets the importance criteria, so we'll go on with the rest of the criteria.

Alex, do you have a question?

DR.CHEN: With a change to the first visit, or reiterative timing?

DR. WINKLER: So, I wonder if we

could have the opportunity to bring it back with it framed as the first visit.

Is that what you all would like? Okay, we'll do it that way.

DR. GLAUBER: Should it be by the first visit, or by a chronological age, for example, one month of age?

DR. PERSAUD: I think it should be chronological. And to me, it doesn't matter where it gets done first. It really should get done first in the hospital, and it's probably -- I think what's worrying me a little bit about this measure is, I think, that we have seen a reduction in SIDS because it's being done as discharge preparation, and it is being done in the office, and I don't know if measuring this is going to get us much further.

I sit also on a fatality team, and we have an unyielding group of deaths that are not going down, and I don't know if this is going to affect that.

There's this group of babies, and I don't know if it's cultural, what it is, that aren't responding to counseling or maybe they're not all SIDS and we don't know what they are. And I don't know if this measure is going to help us close that gap. That's my concern.

DR.CHEN: Right. I don't think you can ever expect to close the gap, because there's a recent article in 2009 in Journal of Pediatrics that pretty much suggested that a certain percentage of kids cannot be modified by whatever we do with a sleep position. And some of them are actually cardiac kids.

DR. GLAUBER: Have you found that ETS exposure is a component of this unyielding group?

DR. PERSAUD: Variably. I mean, we're just not sure at all about them. The pathologist just are completely uncomfortable with calling, what is it, both the cause and the manner of death.

DR. WINKLER: So, does the committee want to give NCQA an opportunity to revise the time frames and bring it back for you to have a second look?

Okay.

DR. MILLER: Is it possible for NCQA to also not only look at the first visit but also look at hospital discharge?

DR. SCHOLLE: At hospital discharge? We'll consider. It's a completely different frame, but we'll take it back and discuss it with some of the other measure developers that are working in that area.

DR. WINKLER: Okay. All righty. Okay. So the next measure for the afternoon is 1381, and this is from the Alabama Medicaid Agency.

Do we have somebody from -- the developer on the line?

DR. MCINTYRE: Yes, this is Dr. Mary McIntyre.

DR. WINKLER: Great. Thank you

for joining us. So, Dr. Glauber, I believe this is yours. This is asthma emergency department visits. This is from work group one.

DR. GLAUBER: This is totally unrelated to any of the other measures in work group one --

DR. WINKLER: Right, right.

DR. GLAUBER: -- so it's sort of a stand-alone.

DR. WINKLER: Yes. We had to vote somewhere, so you caught it.

DR. GLAUBER: Good. I'm glad to have it. It's -- you know, it's interesting that we're only getting to an asthma measure this late in the meeting.

So it is the percentage of patients with asthma one to 21 years of age who have had one or more ER visits during the measurement period.

And just a little bit of background about asthma, in that the current

thinking in the guidelines is that there's really -- it's a heterogeneous disease that should be assessed along two domains that don't necessarily relate or track that closely to each other, one of which is impairment, which is the level of day-to-day symptoms and activity limitations a patient may experience with asthma.

And the other is the risk domain, which is the potential for serious or life-threatening asthma attacks along with impairment in pulmonary function.

So this is really an outcome measure in the risk domain. And the measures that are currently in place and have been endorsed by NQF are really more process measures and in the impairment domain, so this is really the first measure that looks at the risk domain in terms of exacerbations.

And you know, I don't think the importance of this is much in doubt. The developers provide some data from their

Medicaid program that ten percent of their enrollees have been identified as having asthma, and that totally aligns with my patient population in Massachusetts with about a ten percent prevalence.

So again, this is an outcome measure, although in a certain sense, it could viewed as a process measure, and if the timely identification of a child with an ER visit leads to further interventions by primary care physicians to improve their overall asthma management, so as the developers should, they use this as a trigger -- an ER visit as a trigger into a chronic care management intervention.

Is there something you wanted to add there?

DR. MCINTYRE: Well, and I wanted to at least give an understanding on how we came about it, but I'll do that when you finish.

DR. GLAUBER: Okay. In terms of

the scientific acceptability and specifications, the percentage of people with asthma that have an ER visit during a 12-month measurement period, and there's no widespread accepted definition, at least administratively, of how you identify the asthma population.

So in terms of the denominator, what the developers had proposed is a claims-based algorithm which looks at a diagnosis, a visit with an asthma diagnosis, or two -- at least two prescriptions for a short-acting beta-agonist.

And you know, given the age population here where we're included people down to one year of age, there is some potential for mis-diagnosis or misidentification of kids who don't have an asthma diagnosis but who have received a couple of these prescriptions, perhaps for bronchiolitis episodes.

And in terms of validity, the

developers have suggested that identification of a claim for -- with a diagnosis of asthma validly identifies children with asthma.

But they've defined this not based on medical record review of the ER visit, but the fact that of the enrollees who were accepted into this care management program, all of them had asthma.

So you could probably see, this is not a randomly selected population. So you'd imagine that the families of the kids who truly did have asthma would accept enrollment into the care management program.

So I think there are some validity concerns about whether a single claim for an ER visit with an asthma diagnosis truly identifies this population.

There is -- they say that there is potential for stratification by gender and race code, and that's important if there is reliability identification of race/ethnicity because this is an area where there is plenty

of evidence of racial/ethnic disparities in terms of asthma ER visits, hospitalizations, and even mortality.

And it would be important to stratify this population by age, because it's known that children zero to four years of age have a much higher rate of asthma exacerbations in an ER visit.

So if you're doing comparisons of different populations, you'd want to stratify this by age, because if you have a different age mix within the population, that's going to drive differences in the rates.

I don't think there is much concern about validity testing, since it's pretty much a claims-based query.

But I did just want to reemphasize my concerns about the validity given -- also, I think it's good that we're looking at the younger-age population, because this is the highest risk population. But this is also the population in which there's more diagnostic

uncertainty about what constitutes and when you make the asthma diagnosis.

So there could well be variability in provider practices in terms of what they call a wheezing child when they present to the ER. So within one community, physicians may preferentially call these kids bronchiolitis or reactive airway disease, whereas in another community, these kids may be more likely to receive an asthma diagnosis.

So some of the variability in performance may be due not to actual performance but to diagnostic preferences.

But I think, overall, this is an important measure and an important outcome for asthma, and I think fills a need in the asthma measurement space for having an assessment of ER visits.

DR. PERSAUD: Can I ask, we've -- did we do an asthma measure in the child health outcomes group?

DR. WINKLER: I think you looked

at one, but I don't believe you recommended one.

DR. PERSAUD: We didn't -- oh, we did not, okay. Because I remember that the age issue had come up and we discussed, that measure, as it was presented, had every age group.

And I remember we had a discussion about who to exclude, and I didn't remember if it was up to one or up to two is what I'm kind of thinking.

DR. GLAUBER: Yes, and I would also point out that the HEDIS asthma measure excludes kids under five. So we really -- for the majority of people with asthma, the disease starts in early childhood. And we don't have any good measures, so I think this does fill that need.

DR. LIEBERTHAL: I think that Jim gave an excellent summary of the measure. I would disagree with him, however, with the age group.

Children under -- whether it's three, four, or five, NCQA has chosen five to twelve, and almost every group that's looked at asthma measures starts at age five.

Under age five, you get into all of the viral-induced causing of wheezing, and viral trigger is the most common trigger in children under five years old.

Even though there was one fairly recent article that seemed to show that controller medicines did have an impact on these children, many other studies have shown that controller medicines do not have value in these children, and the treatment is oral steroids at the onset of a URI.

I think that the age group under five is a very mixed bag, and so this type of an asthma measure in that group I think would probably not be a good idea.

Also, this is a group where the parents panic pretty easily and will wind up in the ED just from parental anxiety, not

because of the severity of the illness.

The -- so I really would -- I don't like it going down to one year of age.

DR. GLAUBER: Yes, I'm glad you raised that point, because I forgot to mention that I think this is an important outcome measure.

I'm less sold on it as a quality measure for the reasons you said that available treatment can only partially impact this outcome, and children who are receiving the best asthma care will, nonetheless, have exacerbations leading to ER visits.

There's controversy about the effectiveness of preventive measures. And also a good -- when you're looking at the younger population, some percentage of these visits are going to be totally non-preventable because they're going to be an incident case of asthma in a child who has no previous diagnosis.

But given that this is the

population that has the highest risk of ER visits and hospitalizations, I do think it's important to have a measure that's looking at that, even though it's not a performance measure.

DR. MCINTYRE: This is Dr. McIntyre, and I wonder, at what point can I just give you a basic understanding of why we came -- ended up with where we are with this measure?

We actually had the same discussion that you did about the under one, you know, what age did we need to look at.

And basically, I think it would help with the understanding of the intent of the use of the measure.

This actually came about as the result of what we were doing with the transformation grant that we received funding for. And part of that grant was to establish our data-driven quality improvement program, that was specifically looking at putting

together a chronic disease program and having some way to determine whether or not we had actually improved outcomes, and actually ended up with 50 plus members from diverse backgrounds.

And we actually pulled in what we call domain experts, pulmonary pediatric pulmonologists from the university system, as well as pediatric asthma center specialists to talk specifically about the whole issue.

And we did get into, what age do we need to look at, but we determined, because what the intent of this group was to identify patients that could potentially respond from intervention.

And so we were actually trying to really, you know, go beyond, you know, the whole idea about, well, we may end up with somebody that may not actually be asthmatic, that whole discussion came up with using the claims data, and putting in those medications.

But we looked at it as so much

about -- we talked about accountability, and came to agreement that the focus was on the identification of individuals, and not so much on determining whether or not a provider was accountable or performing below a certain measure.

That was the initial beginning of it. So the goal was to improve care and outcome.

So then we actually put the measure into pilot testing in February of 2008, and we called it one of our five asthma missed opportunities, being really careful about what we ended up naming the measure. And it was the asthma emergency department visit.

And the whole goal of this was, we looked at it on an individual, the ability to be able to look at it overall with all of the counties combined, as an individual county measure, but even to allow a provider to be able to click and go down and identify those patients in his practice that had been seen in

an emergency room with a so-called asthma diagnosis, with ultimately the provider being the one that says either they actually have asthma or they don't.

And we have a PCPM population, which is a primary care case management program, so these patients are assigned with a physician.

They may not have even seen them by the time they end up in the emergency room to give them the ability to be able to get those patients in.

And then we combine it with the care management referral to be able to then say, here are patients that have been in the emergency room within this time frame, and what we really want you to do as a care coordinator is work with trying to identify -- you know, we used a CARAT on children. We did an EPA assessment on adults. They did quality of life tools on everybody that was enrolled in the care management program.

And so ultimately, there is an evaluation piece, and the results are that we're using the University of Alabama at Birmingham that we were able to improve the outcomes.

We were also able to move not just the quality of life measures, because that was initiated during intervals within the care management, but also, we were able to actually look at the measures and actually look at where we were a year out, two years out, within the pilot program.

So, you know, part of this being about the claims base, and that whole issue, we basically went with what we could get information with, not having access to state-wide EHR system or HIE with all of the information, and we were able to make a difference for these parents with the children, because it was mostly children in the asthma piece. We did a diabetes one as well.

You know, ended up actually identifying that they had a broader knowledge, a better understanding, they knew what the triggers were, they actually knew what the medications were.

Many of them didn't. They didn't know what an asthma rescue medication was from a controller. They -- we found out there was really a lack of knowledge when it came down to that.

So there were improvements that have been documented that, hopefully, it's in the final stages to get to CMS with UAB planning on publishing the results of the study as part of the pilot program.

So the measure does allow the data to be stratified by race, ethnicity, gender, and geographic area. Apparently, only the county level data has been studied to date. We have actually broken it out now by age and that's not on the website yet.

We're looking at adult versus

pediatric to see if there are differences in the results obtained from that. And the evaluation results are being finalized, with many of the quality of life indicators and the changes identified being statistically significant improvements based on the information that's been done by UAB.

DR. WINKLER: Thank you very much.

I just wanted to correct what I replied to Donna in terms of what we did in the outcome project. Yes, my brain was half asleep.

But indeed, one of the measures, the outcomes activity that you did last year was to recommend, and the measure is minutes away from endorsement, is the asthma admission rate measure, the population measure from AHRQ, which is the admission rate for asthma in children ages two through seventeen per hundred thousand population, so.

DR. JENKINS: I just want to clarify. I think this measure is also just

being proposed at the population level. Is that correct?

And I have a question for the developer, which is whether or not any risk adjustment is necessary to understand variation in this measure, or with a very high level population level, is that not necessary to understand variation?

DR. MCINTYRE: I'm really having a hard time hearing you. You're not clear on my end. And it may be because I'm having to use my cell phone, but the other speaker was clear. Are you near to the mike, or can you hear me?

DR. JENKINS: Yes, two questions. One is just to verify that you're proposing this measure at the population level and not at the provider level.

And two, did you feel that there was any risk adjustment based on severity of asthma or other factors that were necessary to understand variation in the measure?

DR. MCINTYRE: Okay. The population level is what we're proposing it, as far as -- and it's how we actually used it with the pilot program.

But I have to admit that we moved it beyond that at that point. At that point, it's actually become, with some modifications to it, a provider-level.

And it's actually part of our profile report that providers get on a quarterly basis, where they look to see where they are compared with their peer group, not like we did with our population level where we determined targets -- set target goals.

So, it's already moved beyond that even though what we used it with, with CSQ, is population level.

Risk adjustment, as far as with severity, we worked with -- we didn't do any risk adjustment with this, because what we were trying to do is identify people and allow providers to be able to identify those people,

children and adults, who have been to the emergency room with a --

(Temporary failure of telephone connection.)

CO-CHAIR MCINERNY: Any further discussion?

MS. WINKLER: I just wanted to go back, because Donna reminded me of the previous measure that this group put forward, in terms of the age issue.

This measure, as presented, Jim, do you agree? I mean, it's children just less than age 21, right? Okay, I wasn't sure. It's one to 21.

And the previously recommended measure was two to seventeen. You know, this is a perfect example of the need for harmonization, folks, I mean.

DR. MCINTYRE: I'm sorry. I'm also at the management conference, and I'm trying to do two things, so I am so sorry.

MS. WINKLER: No problem. Okay.

So we were just talking about age.

DR. MCINTYRE: Yes.

MS. WINKLER: I'm asking the committee, in terms of other measures that we've endorsed around asthma, particularly asthma hospitalization, the age range is two to seventeen, and this is one to twenty-one. We really don't want them all over the board like that.

DR. GLAUBER: Well, I think if we went with age two, that would address some of the mis-classification potential around how the denominator is including kids who would qualify just on the basis of having two Albuterol prescriptions.

DR. MCINTYRE: And this is Dr. McIntyre again. We're fine with the two, we just put it in because our experts wanted to look for younger children.

But let me ask something else. And I know about your 17, which is one of the issues that we're having with some of the

HEDIS measures, with the Medicaid population, they are covered, as long as they're covered, up to the age of through 20, okay?

So that when we drop them before 20, you know, when we have measures and we don't have anything to go with that population, it's almost like they've dropped off a cliff and they're in no-man's land for the Medicaid.

So that's the thing that we have with even some of the measures that you all have with not modifying them. Okay, I just thought I needed to say that.

DR. JENKINS: Also, I didn't hear the answer to the question, because you had dropped off, about risk adjustment?

DR. MCINTYRE: Oh, the risk adjustment was that we didn't find that it was necessary to be done as part of this because of what we were using the measure for, but we did do it with AAHRQ and their asthma return on investment calculators, we put in the

information in the pilot.

And basically, when you look at severity, we were trying to -- you know, it was almost like it was like -- when you started getting two or more ER visits.

We actually looked at the number of visits, even though this just picks up anybody with one or more, because we wanted to try to get those children and intervene, and adults, okay?

But we didn't do any other risk adjustment. But we have done some testing beyond what we've done.

DR. GLAUBER: I would also argue against risk adjustment. There was a -- just a study presented, it's not published yet, but it was presented at CHSS meeting from Connecticut looking at a series of their asthma hospitalizations.

And they reported that roughly half of their admissions were in kids previously thought to have

intermittent asthma or mild, persistent asthma. And in terms of their ICU admissions, that having being thought to have intermittent asthma in the past significantly increased -- no, it was much more of an association with an ICU admission.

So, you know, in terms of exacerbation risk, it doesn't seem to -- as I said, it doesn't necessarily tie very closely to disease severity in terms of the ways in which we currently stratify for disease severity.

So, you know, basically, a kid with very mild asthma can have a pretty serious or life-threatening exacerbation.

DR. SCHWALENSTOCKER: This is Ellen, and I realize this is a population-based measure, but one worry I have about it is the unintended consequence of taking kids out of receiving care, even if it's in a sub-

optimal setting. So I wondered if you --

MS. WINKLER: Ellen, we're having trouble hearing you. You're fading in and out.

DR. SCHWALENSTOCKER: Is that better?

MS. WINKLER: Yes.

DR. SCHWALENSTOCKER: Can you hear me?

MS. WINKLER: Yes, that's better.

DR. SCHWALENSTOCKER: So, one question I had, I realize this is a population-based measure.

But I wondered if you had any experience or evidence that the measure itself resulted in discouraging seeking care for children, even if it's in a setting we might not prefer.

DR. MCINTYRE: No, it actually did the -- it worked in the opposite direction. One of the things that we identified, is many of these children had not even been seen by

their primary care providers, and not just children, adults, that they didn't have an office visit, but they had an emergency room visit.

And what we ended up doing with this is we were able to get them and connect them to their primary care, their medical home. So it actually worked in the opposite direction with getting them into care, which is one of the goals of the program as well, okay, to connect them to a medical home.

MS. WINKLER: Okay, thank you. Okay, any other discussion?

All right. So how does the committee feel that this measure meets the importance criteria? Yes, or no? All say yes?

Nos? One.

Marlene and Ellen?

DR. MILLER: Yes.

DR. SCHWALENSTOCKER: Yes.

MS. WINKLER: Okay. All right,

that's 14 yes, 1 no.

All right. Scientific acceptability of the measure properties. How many believe it meets the criteria completely?

DR. GLAUBER: Is that with the amended definition of the measure to start at age two? Is that where --

MS. WINKLER: I think that would be with the recommendation for age two.

Completely?

Partially?

Minimally?

Ellen and Marlene?

DR. MILLER: Partial.

DR. SCHWALENSTOCKER: Partial.

MS. WINKLER: Okay. All right. Usability, completely meets? Two.

Partially meets?

Minimally?

Marlene and Ellen?

MS. MILLER: Minimal.

DR. SCHWALENSTOCKER: I would say

partially.

MS. WINKLER: Okay. Feasibility, completely meets?

Partially?

Minimally?

Marlene, Ellen?

MS. MILLER: Minimal.

DR. SCHWALENSTOCKER: Partial.

MS. WINKLER: Ellen?

DR. SCHWALENSTOCKER: Partial.

MS. WINKLER: Okay. Questions, discussion? No?

Okay, recommendation for endorsement, all those yes?

Any no's?

And Marlene and Ellen?

DR. MILLER: No.

DR. SCHWALENSTOCKER: I vote yes with the age difference.

MS. WINKLER: Okay. All right. There's 13 yes, 3 no. All right.

So, moving on down the last, the

next measure is, we go to two measures of vision screening, which is more the area that work group one was functioning in.

And the first one is 1412, preschool vision screening in the medical home. This is brought to us from the American Academy of Pediatrics.

Do we have somebody from the measure developer on the line? Okay. Are we expecting somebody? Oh, we are. Are we early? Well, we're not early.

Is anybody from AAP on the line? No.

Well, why don't we try and do the NCQA measure first, and maybe the other folks will show up.

So it's 1398, vision screening.

Kathy, I think that's yours.

This is, again, this is another measure that's split into three ages, but it's vision screening by six years, vision screening by thirteen years, or vision

screening by eighteen years.

DR. JENKINS: Sorry, I think this is similar to some of the other measures that we've seen from NCQA, where they've put together recommended vision screening, that there are three age groups, as we have just heard within their comprehensive well child care.

And -- just looking for the numerator statement, I just have to scroll down through all their evidence.

There are some recommendations from the TAP work group here which has some concerns about the screening methodologies and the specifics of the actual types of screenings, and a number of concerns about low sensitivity related to the specific instruments that were being used, and a real call that any screening recommendations only use high sensitivity instruments that are more robust.

And I would just call the group to

look at the comments from the TAP work group. I think, Ellen, you had been on the calls. We talked about this earlier, and I might ask you to comment on what the TAP group had suggested for both of these measures.

DR. LIEBERTHAL: The Tap Work Group included an ophthalmologist and an optometrist, and I was a little bit surprised by their comments regarding the sensitivity and specificity of the standard tests used in pediatric offices, which are generally the wall charts, whether they be the Snellen symbols, the E -- what are -- they call it the falling E or something like that, or the letters.

And they said that those are very low sensitivity. They really didn't -- they recommended one type of test that's done at five feet using different wall charts and saying that they're much more accurate because the child is closer to them, you have the child more engaged, and you don't have that

twenty-foot space of distraction.

But beyond that, the vision screening devices, none of them appear to be adequate, really good or practical.

So what I came out of that meeting with was, you know, there are no really highly sensitive tests available in the general provider office. However, the -- it's a lot better than nothing.

The division into the three groups, their feeling was that getting the child between age four and six was the prime group, earlier if possible, but that's not always the case. And doing the testing for the twelve-year-olds and the eighteen-year-olds didn't seem to be as important.

So their leaning was towards just the one age group, and I guess it would be nice if we could optimize the test that was used.

And they said that no specific method was mentioned in the measure, but I'm

not sure there's one available that would be appropriate.

DR. JENKINS: Yes, and just to reiterate that a lot of the importance criteria for this and for the other measure that from the AAP related to vision screening were about the risk of missing and intervening on amblyopia.

And that there's a time issue there that's a little bit disconnected from the six, thirteen, and eighteen age windows that we're looking at here.

In terms of what was actually specified as the numerator details, it's a similar documentation criteria that we've seen in some of the other NCQA measures, and it's for -- there is some specification about distance visual acuity, specifically in each eye. That was important.

And I guess in this one, the evidence of confirmatory testing or referral or a follow-up visit was included. Is that

correct?

So in this case, you did include the referral criteria, as opposed to the others? So I guess that's a difference.

The denominators are very similar to what we've seen before with the two year look back and the birth date and the requirement for a visit with a primary care provider in the prior 12 months.

There wasn't any risk adjustment for the process measure. I think that's really about it.

I guess my personal concerns are similar to some of what we've talked about previously, but the major issue that was brought up was really what Allan alluded to from the TAP, and our general concerns about recommending or embedding into quality metrics or performance metrics screening tools and testing that are not widely regarded by experts as of value. And that introduces a problem.

DR. LIEBERTHAL: This was the first time I had heard that the standard ways of testing have such low sensitivity. And it came from two individuals who received the referrals, and I'm sure that they were quoting evidence-based information.

If we throw away the testing that we traditionally do because it's not accurate enough, we have two choices.

We either substitute something that's more accurate, and from what I could gather from these two individuals, the only thing that that would be at the current time would be referral to a specialist, which some people have advocated that all children should see an optometrist before they start school, or do nothing. And I think the referring all children to an optometrist is a logistically and cost-wise very problematic.

Doing nothing is, I think, even though it's not as good a test as we would like, it's still better than doing nothing.

CO-CHAIR MCINERNY: Yes, Allan, did they say that the problem was they were getting too many children referred who failed a test and had normal vision, or did they say -- and I always forget sensitivity and specificity, which is which. Did they say we were missing kids --

DR. LIEBERTHAL: Missing kids. Sensitivity would be a high false negative rate.

DR. JENKINS: Just to reiterate the point, I think it might be on the TAP, because the TAP said most children who fail vision screenings never receive care, which was another issue.

So it was also wanting to kind of close the loop, not just on the screening, but also on the rest of it.

And there's a comment here from the TAP work group from the AHRQ national Advisory Council on Health Care Research and Quality Subcommittee on Children's Health Care

Quality Measures voted the vision screening measure off the core set of 25 measures that the committee recommended to the secretary for Medicaid and the CHIP programs.

Chair Rita Mangione-Smith said it was not a reflection of the importance of vision or eye health among children, but rather the opposite, a consensus among the Subcommittee that current vision screening is, as reported, fragmented, disorganized, and unaccountable to the desired protected child health outcomes. In her words, vision screening is not ready for prime time.

So, that's the issue -- that's the history here, and I believe that some of their TAP comments were simply reflective of a similar perspective.

DR. GLAUBER: And I recall that there was a TAP comment saying that the -- you know, the risk of the low sensitivity is that families would receive false assurance that their child had normal vision and then maybe

not be as responsive to any ongoing problems that may emerge.

DR. JENKINS: Exactly. So I guess we need the measure developer to address some of these issues. And also, for me, the time frame issue, where there's a disconnect between some of the early detection importance to preserve vision as well as get kids into treatment and services, and the level of the measure in the older age groups.

MS. SCHOLLE: So, we actually tested a different -- a more complicated measure than what we presented here. We changed some of the numerator requirements based on the results, because to try to respond to the concerns about what's most important for children at age six and age thirteen and eighteen, our panel felt that it was important to do the screening for both the younger kids and the teenagers because of the changes in vision that happen during adolescence.

I can't speak to the issue about the sensitivity and specificity. We originally tested specs that were based on the specialist recommendations for screening, and we found that those were unworkable.

Just, we couldn't document what was happening that way in the primary care setting in the medical records.

I don't know about the sensitivity and specificity of these screening approaches. I'm not sure that -- I don't remember which measure was used -- was proposed and reviewed and is the point of the discussion about the CHIPRA core measures, and what you've quoted from that report, but it was not this measure.

And I thought that it actually had to do with vision and hearing screening that was happening in the schools and whether it was being recorded somewhere else.

So, anyway, I can't refute -- the piece that I don't know about is whether this method of screening that's done routinely or

in primary care practice is sufficient for identifying vision problems in children.

DR. LIEBERTHAL: I think what they're saying is that it is not sensitive enough to be the standard of -- for screening.

And I personally -- I'm not quite sure how to deal with that, because I cannot imagine not doing -- you know, making some effort to determine a child's vision.

DR. GLAUBER: Is the sensitivity problem particular to the younger age group, or is it the older age groups as well?

DR. LIEBERTHAL: They can find their discussion, and you'll see what the next -- the other measure, the AAP measure. They limited their discussion to the under-six-year-old.

They felt that the measuring for the thirteen-year-old and the eighteen-year-old was unimportant -- did not have a level of importance.

I think probably -- you know, I'm

not -- I think it would probably be more reliable, saying, isn't that what an optometrist does when you first walk into their office, is they throw the letters up on the wall and say, can you read those?

CO-CHAIR MCINERNY: Now, this is an interesting problem because one of the -- the first study that the Pediatric Research in the Office Setting folks did, and this is now close to 20 or maybe even 25 years ago, was how often pediatricians were doing vision screenings at three and four years of age.

And it turned out they weren't doing them very often. And now, I'm saying, well, maybe they were right, because they weren't that good.

(Laughter.)

Although, you know, this is sort of an article of faith for pediatricians in the Academy of Pediatrics is to do the vision screenings. And certainly, in the three, four, and five years old, four to try to

prevent amblyopia.

The other issue I had, with the older children, is, at least in New York state, and I don't know if this is true in other states, maybe you can answer that, but the nurses at school do vision screenings for kids once they're in kindergarten and all the way through high school at some rate.

And so that becomes duplicative, if they're doing it and we're doing it, then why do that?

MS. BERGREN: That's true. It varies from state to state what the number of times a child is screened. But it's pretty common in the younger ages, in the early grades, of kindergarten first, and then beyond that, it's every couple of years. But it goes away in -- but not every state has any requirements in the high school ages.

DR. LIEBERTHAL: In California, it's very spotty. It can vary from school district to school district, and with a large

school district such as Los Angeles, from school to school, because they don't have the resources.

CO-CHAIR MCINERNY: It seems like you need a proposition on that.

(Laughter.)

DR. LIEBERTHAL: Well, if you're going to have one proposition, you have to two contradictory propositions. You've got to be able to read it.

(Laughter.)

MS. BERGREN: I guess I'm concerned that there's nothing to substitute for this, and that it doesn't call for a particular type of screening instrument in the criteria.

And to not recommend this would -- I don't see it being substituted by recommending that all children get an ophthalmologist or an optometrist visit. I see this as reducing the number of kids that are going to be seen, if we don't agree with

it, as opposed to increasing it.

MS. WINKLER: I listened to the TAP call also, and the TAP members were absolutely in this quandary.

I mean, I think they wanted to make people aware that the traditional vision screening tools aren't as good as you think they might be, and that awareness was an important factor, but that they certainly were not advocating doing nothing. That was not a good alternative either.

But the fact that there probably is a lack of awareness that the most commonly used tools aren't particularly sensitive is something that needs to be more widely understood.

CO-CHAIR MCINERNY: Well, is this another perfect is the enemy of good enough situation here?

DR. GLAUBER: Well, I just wonder, since the TAP said the importance was a lot less for the thirteen and eighteen-year-olds

and given that there is sort of a paucity of measures and recommendations, relatively speaking, for the six-year-old population, whether we might amend the measure to just focus on the six-year-old, and that would set up the challenges for trying to figure out the most feasible and sensitive strategy for screening.

MS. WINKLER: You don't need to amend the measure. These are three separate measures.

DR. LIEBERTHAL: And measure 1398 is very similar, but only deals with the child under six years old.

MS. SCHOLLE: Would it make sense for us, since I was not able to listen to the TAP call, I wonder if it would make sense for us to talk with some members of that group to see if we could get some advice on refining the numerators?

MS. WINKLER: A couple of options. Certainly you can. We also posted the

recording of the TAP call on our website, so it's available for anybody to listen to.

CO-CHAIR MCINERNY: Okay, well, I think we've had enough discussion. Let's take a vote.

MS. WINKLER: All right. I heard a sense that you would like to focus in on the measure for the six-year-olds. And not so much on the thirteen to eighteens. We'll split those.

And we're talking about the measure as is, though I think it's -- NCQA sounds, you know, open to hearing the feedback and seeing what they might be able to do to address some of these concerns, but we don't know what that's going to be.

Is this another measure you want to table, or do you want to go ahead and vote on it now? Come on, gang.

DR. LIEBERTHAL: I think we can vote on it.

MS. WINKLER: Vote on it now?

Okay. So we were just talking about the age six measure, and how many feel that it meets the importance criteria?

Marlene and Ellen? Marlene and Ellen, are you still with us?

DR. SCHWALENSTOCKER: Sorry, I was on the -- yes.

MS. WINKLER: Marlene, are you still there?

It looks like we lost her too. Okay.

Then the scientific acceptability of the measure, and we're talking about the six-year-olds, okay?

How many feel it meets the criteria completely?

Partially?

Minimally?

Not at all?

Are there any no's?

And Ellen?

DR. SCHWALENSTOCKER: Minimally.

MS. WINKLER: Okay, thank you. Usability, completely meets criteria?

Partially meets criteria?

Minimally meets criteria?

And Ellen?

DR. SCHWALENSTOCKER: Partially.

MS. WINKLER: Okay. Feasibility, completely meets? Partially? Okay.

Minimally? All right.

Ellen?

DR. SCHWALENSTOCKER: Partially.

MS. WINKLER: Okay, great. All right. So, recommend for endorsement? Just the one, just the six-year-old. Yes?

Nos? One.

Ellen?

DR. SCHWALENSTOCKER: Yes.

MS. WINKLER: Okay, twelve yes, one no.

Okay. Do we have anybody from the American Academy of Pediatrics, measure developers for measure 1412? Anybody out

there?

DR. LIEBERTHAL: I don't think they're on. I actually -- before I even knew that this was going to be presented at this meeting, the prime author, who's chair of the section of ophthalmology in the AAP called me for advice on developing a measure. So I have a little background on the measure, so I think in that case, maybe I can present that and not vote.

CO-CHAIR MCINERNY: So Kathy already left, so 1398 is a -- 1412, I'm sorry, I got them mixed up.

This is presented by the AAP. The background on it and the importance of it are pretty much the same as the NCQA measure under six years of age.

This is number of preschool children under five years old that receive visual acuity testing or photo screening in the medical home is the numerator.

It's a pretty straightforward, and

each has a CPT code.

By the way, and with regard to the TAP comments, there are no CPT codes that differentiate the type of screening done, other than whether it is screening test of visual acuity or the photo screening, which involves devices that are used.

So, even if they came up with other methods of doing visual acuity screening, there's currently no way of identifying what is done.

The denominator, again, is a very simple denominator, all children under five years old who attend a routine well child visit in their medical home.

He was -- when questioned about, does it have to be in the medical home, what about kids who are already seeing an optometrist or an ophthalmologist, the author wanted to maintain that it be done in the medical home whether or not they were being seen outside, as just something that we should

do on a routine basis.

And exclusions are basically that you can't do the screening due to the patient being unstable or uncooperative or the parents have refused screening. And I don't know how you would extract that from the medical record on a reliable basis, but those are the only exclusions.

This measure has not been tested for validity or in any other -- or reliability.

The comments of the TAP were essentially the same. They preferred this measure, first of all, because they think that screening by five years is more -- is better than waiting until six years, and also they were -- they thought that the first measure, the previous measure, was an all or none as far as the age groups, so their leaning was to approve this measure as opposed to the other one.

But the comments on the numerator

are pretty much the same, the accuracy. The question was, on the TAP, just quoting, children with any vision-related symptoms should be excluded from screening, as children with vision-related symptoms should never receive vision screening of any type, and should be tracked directly into care by an optometrist or pediatric ophthalmologist, so they were discounting the exclusions as being an indication to -- for referral.

DR. PERSAUD: And I noticed, this one had the USPSTF, strength of evidence is a grade B, and I guess that's why I voted the other one down, because I scrolled ahead and looked at this one and thought it was a better measure.

DR. LIEBERTHAL: I didn't look at the other one about the USPSTF -- but were they looking at age group, or were they looking at all three together?

I really didn't look at that closely. But they should be the same, because

you're basically doing the same thing.

DR. PERSAUD: I think it had to do with the age cutoff. What the USPSTF said is that for older children, you don't make any change in visual acuity, but for younger children you can help them with amblyopia and blindness.

MS. WINKLER: These are two measures that are essentially very, very similar, and what we do when we look at similar measures is we evaluate each of them against the criteria and compare.

If you feel that one, by the way, it's specified or the way the data source is or whatever, look at how they meet the criteria.

If one clearly meets the criteria better, then there will be a preference for that measure.

If they're the same, then at some point, we'll need to make a decision, because we would not really want to see two measures

this much alike being recommended. There would be no purpose in that.

So I do want you to take a look at it from that perspective.

CO-CHAIR MCINERNY: So, in that situation, should we go through the regular voting on the measure, and then at the end of that, take a vote as to whether we prefer this one or 1398?

MS. WINKLER: Yes.

CO-CHAIR MCINERNY: Okay. Sounds good.

MS. WINKLER: Do we feel like everybody's had a chance to look at the specifications of both measures, in particular, of this measure as we do it? Sarah?

MS. BROWN: Can you just tell me quickly what the difference between the two of them is?

MS. WINKLER: Basically age.

MS. BROWN: That's all? And

medical --

MS. WINKLER: And specifying medical home rather than just the --

DR. LIEBERTHAL: And the exclusions. This one does not specify follow-up.

DR. SCHWALENSTOCKER: This is Ellen. I mean, you then have to have a measure of medical home to go with this?

MS. WINKLER: Ellen, we can't hear you.

DR. SCHWALENSTOCKER: I'm sorry. My question is about the medical home, and how that's being defined, and if it's adequately defined in the measure.

MS. WINKLER: I think she's asking about the specifications for the term medical home and whether that would be commonly understood and commonly applied.

DR. GLAUBER: It seems like it's in the definition but not really in the specifications, so, I don't know if that

really has an impact.

MS. WINKLER: Okay.

DR. LIEBERTHAL: I have concerns about that also because if a child is being seen by an ophthalmologist or an optometrist, that trumps the vision screening in the medical home.

However, when I presented that to the AAP and the author, they wanted to keep it in.

MS. WINKLER: Okay. So, on this measure, 1412, how many feel that it meets the importance criteria? So I would assume they would be the same, right? Okay.

Any no votes?

Oh, Allan, you're abstaining.

Ellen, what were you?

DR. SCHWALENSTOCKER: Yes.

MS. WINKLER: Okay, great. So the scientific acceptability of this measure. Does it completely meet criteria?

Partially meet criteria? Nine.

Minimally meet criteria? Two.

Ellen?

DR. SCHWALENSTOCKER: Partially.

MS. WINKLER: Okay. Usability criteria, completely meet?

Partially meet?

Minimally meet?

Ellen?

DR. SCHWALENSTOCKER: Partial.

MS. WINKLER: Thank you.

Feasibility, completely?

Partially? All right.

Minimally?

Ellen?

DR. SCHWALENSTOCKER: Partial.

MS. WINKLER: Okay. And so how many would recommend this measure for endorsement? And it would be a time-limited endorsement because of the lack of testing.

Yes?

Any no votes?

Ellen?

DR. SCHWALENSTOCKER: Yes.

MS. WINKLER: Okay. Now, in terms of, you looked at two measures that are very similar. In terms of your evaluations of the rating of the criteria, they're remarkably similar.

Generally, you felt the scientific acceptability of either measure was -- more votes for the minimal usability, partial, feasibility, partial.

I am concerned about, the scientific acceptability of the second measure was rated more at partial, where it was minimal on the first one.

Jim, you have a question?

DR. GLAUBER: What may differentiate these measures in terms of harmonization is the perceived importance of the target age ranges, and we only got to vote on that as a yes/no.

MS. WINKLER: Okay.

DR. LIEBERTHAL: In determining

the measure that NQF endorses, is there a mechanism for harmonizing the measures, taking the age from one and the wording from another?

MS. WINKLER: Not easily. Let me put it to you that way. I think there's a general concerted effort to want to achieve harmonization. It's much more easily said than done, and the history of the measures have their own reasons for doing things the way they do, and that does not make them very amenable to change.

NQF had a recent harmonization work group that looked at the whole issue around harmonization, and they felt that the harmonizing at the end of the process has very limited success, and that harmonization is best achieved at the early stages of development of measures.

So, and we've certainly found that to be our experience, so this one becomes a difficult one.

I guess from the perspective of

the committee in terms of these two, what do you see to be the real strengths of one over the other? Do you have a preference of one over the other, and why?

DR. CHEN: Can I just raise a question about the difference between medical home and non-medical home setting?

Is there any reason why we should prefer the medical home setting versus the non-medical home setting? I mean, I think age makes a difference for me, but aside from that, it's really the main difference is the medical home versus the non-medical-home setting, right?

CO-CHAIR MCINERNY: Well, I think the reason for that is the reason for the Academy of Pediatrics' stress on the medical home is because that the primary care physician in the medical home is to be responsible for the patient's health in the broadest sense of the term, and that therefore, it would be important that this be

done in the medical home.

MS. BERGREN: Were you finished? For early childhood screening that's done in the school setting, there are highly trained screeners who do that screening, and that is one of the only things that they do. They don't really have a lot of other responsibilities.

So, from that respect, I don't think there would be that much difference. You know, the screening that's done in the early childhood setting is usually pretty good.

DR. ZIMA: I guess I had a preference for NCQA just because the specifications were better operationalized.

And then the other question I have is, and it's really just question is, if this is close, does membership to AAP influence decisions among our group?

I'm just saying, but if we were to look at -- I'm just thinking in terms of your

position, having to defend the choice, I think it would be something simply to explore as a sensitivity analysis, just to make sure.

CO-CHAIR MCINERNY: Well, a couple of things. One, not all children are in an early program, educational program, so you'd miss a bunch, and that's where the medical home would be more important.

Number two, I would certainly, if folks are more comfortable, I think those of us who are AAP members could recuse ourselves from the vote.

DR. QUIRK: We've already established that these screenings got poor sensitivity, so maybe it's less -- probably there isn't a statistically significant difference in who does the screen, whether it's a school nurse or a technologist or a pediatrician or a family doctor or an optometrist, so who cares?

I think timing is very important. I think that the kid should have a visual --

it doesn't get everything, but if you can't read the letters on the page or on the blackboard, you're going to be stigmatized in about five milliseconds in kindergarten.

I have a very vivid memory from 1951 where we were all screened by the school. We had already decided who in the kindergarten class couldn't see very well, and we kind of ostracized them a little bit.

So I think it's very important, before you get into an organized, socialized, or pre-socialized activity, that that's taken care of, and then that there's a referral.

I don't care if it's in a home, I don't care if it's in an office or a school nurse's office.

DR. LIEBERTHAL: In response to the question, I think it was said at the beginning of the meeting that we're here for our expertise, not representing any particular organization, and I take that very seriously.

The only reason I recused myself

is I had background information from the developer which I thought disqualified me from voting, but it wasn't because I'm a member of the AAP or was appointed for this by the AAP.

DR. ZIMA: And for the record, I don't think AAP members should recuse from the vote. I'm just anticipating a close vote.

MS. WINKLER: Well, how about this as a first step. Why don't we take a bit of a straw vote, and just see.

With -- we look at the differences between the measures that are primarily around age, primarily the specification around the medical home, and the exclusions. Did I miss anything, are those --

DR. LIEBERTHAL: Yes, the NCQA included follow-up --

MS. WINKLER: Oh, okay.

DR. LIEBERTHAL: Which the other measure did not.

MS. WINKLER: Okay, the follow-up on the NCQA measure. All right. So those are

the sort of the main issues.

So we have some of the folks from AAP on the phone? Hello? Hello?

Yes, hello?

DR. PLUMMER: Hi, can you hear me?

MS. WINKLER: Just barely.

DR. PLUMMER: Okay. This is Janelle Plummer from the AAP.

MS. WINKLER: Okay. Great. We've been talking about your vision screening measure here for a while.

DR. PLUMMER: Yes, I heard.

MS. WINKLER: Oh, good. All right. Did you have any -- if you've been listening, then, do you have any comments in response to some of the discussion?

DR. PLUMMER: I just wanted to be clear, the two measures are NCQA and AAP, and the issues are around the setting and the inclusion of follow-up.

MS. WINKLER: Yes.

DR. PLUMMER: Are there any other

concerns?

MS. WINKLER: Age was the discussion point.

MS. DR. PLUMMERER: Okay. Okay.

MS. WINKLER: Did anybody have a question you wanted to ask the measure developer on this measure?

Okay. Dr. Lieberthal kind of represented the measure. Were there any questions the developer wanted to ask of the committee?

DR. PLUMMER: No.

MS. WINKLER: Okay. All right. So, let's get a sense of -- we don't have the full committee here.

We had a bit of an attrition here, so maybe this will be sort of a first round, see how it goes sort of thing, in terms of, we've talked about the differences between the two measures.

So how many of you would prefer -- having to make a choice, how many of you would

prefer the measure from NCQA, at age six?

And how many would prefer the measure from AAP, the medical home, blah blah blah?

Is there anybody who didn't vote? Allan recused himself.

DR. SCHWALENSTOCKER: Yes, I'm going to abstain, because I think I missed something on the NCQA measure and I don't want to state a preference.

MS. WINKLER: Okay. At this point, you've indicated a preference for the AAP measure. Was there something that -- was there any particular follow-up with that, any clarifications, any issues you wanted to further raise with the AAP?

DR. LIEBERTHAL: Janelle, you're still on the phone?

DR. PLUMMER: Yes.

MS. LIEBERTHAL: Yes, It's Al. The only thing I liked about the NCQA better than the AAP measure was that it specified

action to take if the vision screening test is abnormal, and I think that is something that, if the authors would consider, might be added.

DR. PLUMMER: Okay.

CO-CHAIR MCINERNY: The only problem is, I don't believe, as I look at the numerator, that in the NCQA measure, it didn't say follow-up in the numerator. I'm trying to find it again.

DR. LIEBERTHAL: In 2A.3, numerator details --

CO-CHAIR MCINERNY: Oh.

DR. LIEBERTHAL: -- it says, documentation must include the data note indicating the following, and there were three bullet points.

CO-CHAIR MCINERNY: Okay. It's not in 2A.1 but 2A.3. Okay. Very important. Thank you.

Yes, and I agree, I think the follow-up would be important if we could include that in the AAP measure.

MS. WINKLER: Okay. All right.

DR. PLUMMER: Would there be an opportunity for modification, prior to the endorsement?

MS. WINKLER: Yes, I think there's an opportunity. These are recommendations to you that you could bring back to the committee, but we'd have to do that in a reasonably short period of time.

DR. PLUMMER: Okay.

DR. RAO: Just a quick question. That was a time-limited measure, the AAP?

MS. WINKLER: Yes.

DR. CHEN: Only the AAP is. I think that's part of the reason why I prefer it, too.

CO-CHAIR MCINERNY: We've done very well. We actually got through all of today's measures.

However, we did have four measures left over from Monday, from yesterday, that we did not get to, and I don't think we have time

to go into those now.

We're supposed to -- many of us are going to be leaving shortly. We do need time for the public comment, and then the wrap-up.

MS. WINKLER: I think we've pushed you fairly hard today, and I see many signs of fatigue around the table, and I think that would be reasonable.

We have several conference calls already established for follow-up, so I think we'll be able to discuss those measures at that time.

So, I think at this point, operator, is there anybody on the phone?

OPERATOR: Not at this time.

MS. WINKLER: Thank you.

Is there anybody in the room who wants to say something?

All right, so public comment was relatively short. It's been --

CO-CHAIR MCINERNY: Well, then

we'll do another measure.

MS. WINKLER: Yes, really? I mean, it's been a really long two days, and we realize that this was a very ambitious agenda. You guys have really done a remarkable job of getting through as much of it as you did, and thank you all for hanging in there and pushing yourselves as hard as you did.

So, what we're going to do is do a quick summary tomorrow of the decisions that you made, and we will kind of circulate them so you have an interim view of the work that we did.

But we still have measures to go, and you will be hearing from us in the next couple of weeks in terms of the agendas for those follow-up conference calls.

Suzanne, did you want to talk about any of the --

MS. THEBERGE: Nope, just we'll be in touch with you with the call-in information, the agendas. The dates are up on

the slide set. Let me know if you have any questions.

If you know now that you are not able to attend one of those calls, and you haven't told me already, please send me an email and let me know, especially if we need to cover one of your measures that we missed yesterday, I'll make sure to get that on a call when you're available. And I just want to say thank you to everyone.

CO-CHAIR MCINERNY: I did a first rough count, of the 19 measures that we considered yesterday, we voted yes for 10 of those.

And of the seventeen measures that we considered today, we voted yes for nine of those. So we're running about -- batting 500 on those.

MS. WINKLER: There's no normal -- it's very variable.

Yes, and I think that's why I wouldn't want to push them.

CO-CHAIR MCINERNY: Right.

MS. WINKLER: I think everybody's -- I think fatigue is setting in.

CO-CHAIR MCINERNY: Great. Thank you, everybody.

(Whereupon, the above-entitled matter was concluded at 2:33 p.m.)

1

2

3