NATIONAL QUALITY FORUM

+ + + + +

PULMONARY AND CRITICAL CARE STANDING COMMITTEE

+ + + + +

TUESDAY
MARCH 15, 2016

+ + + + +

The Committee met at the National Quality Forum, 9th Floor Conference Room, 1030 15th Street, N.W., Washington, D.C., at 9:00 a.m., Dale Bratzler and David Lang, Co-Chairs, presiding.

PRESENT:

DALE BRATZLER, DO, MPH, Chief Quality Officer, OU Physicians - Oklahoma University Health Sciences Center, Co-Chair

DAVID LANG, MD, Chair, Department of Allergy and Clinical Immunology, Respiratory Institute, Cleveland Clinic, Co-Chair

GERENE BAULDOFF, PhD, RN, FAAN, Professor of Clinical Nursing, The Ohio State University College of Nursing

KENNETH BENSON, Board Member, US COPD Coalition

CURTIS COLLINS, PharmD, MS, Specialty Pharmacist, Infectious Diseases, St. Joseph Mercy Health System

BRUNO DiGIOVINE, MD, MPH, Division Head, Pulmonary, Critical Care and Sleep Medicine, Henry Ford Medication Group

TODD DORMAN, MD, FCCM, Professor and Vice Chair for Critical Care Services, The Johns Hopkins Hospital

KIM ELLIOTT, Q, PhD, CPH, Administrator, Clinical Quality Management, AHCCCS

WILLIAM BRENDLE GLOMB, MD, FCCP, FAAP, Senior Medical Director, Superior HealthPlan

STEPHEN GROSSBART, PhD, Senior Vice President and Chief Quality Officer, Mercy Health
JAMES "MITCH" HARRIS, PhD, Director, Research & Statistics, Children's Hospital Association *
EDGAR JIMENEZ, MD, FCCM, Vice President, Critical Care Integration and eICU - System, Baylor Scott & White Health
ELLA KAZEROONI, MD, MS, Professor, Department of Radiology Cardiothoracic Radiology Division, University of Michigan Health System *
THOMAS LAMPONE, MD, Senior Medical Director, Florida Blue
RICHARD MURRAY, MD, Vice President and Deputy Chief Medical Officer, Merck and Co., Inc.
JAMES O'BRIEN, MD, MS, VP, Quality and Patient Safety, OhioHealth Riverside Methodist Hospital
PATRICIA J. OHTAKE, PT, PhD, Associate Professor, University of Buffalo
SUSAN POLLART, MD, Associate Tenured Professor; Senior Associate Dean for Faculty Affairs and Faculty Development, University of Virginia, Department of Family Medicine *
CRYSTAL RILEY, PharmD, MHA, MBA, CPHQ, CHPIT, Manager, Healthcare Policy, Baxter Healthcare
CHRISTINE SCHINDLER, PhD, RN, CPNP-AC/PC, WCC, Clinical Assistant Professor of Pediatrics, Marquette University College of Nursing
DAVID STOCKWELL, MD, MBA, Associate Professor of Pediatrics, Children's National Medical Center *
CHANA WEST, RN, MSN, Associate, Booz Allen Hamilton

NQF STAFF:

HELEN BURSTIN, MD, MPH, Chief Scientific Officer
ANN HAMMERSMITH, JD, General Counsel
ELISA MUNTHALI, Vice President, Quality Measurement
JANINE AMIRAULT, Project Analyst
POONAM BAL, Project Manager
SHACONNA GORHAM, Senior Project Manager
ROBYN Y. NISHIMI, PhD, Consultant

ALSO PRESENT:
MARY BARTON, MD, MPP, National Committee for Quality Assurance
CHARLIE BRUETMAN, MD, The Lewin Group
SUE FRECHETTE, BSN, MBA, American Thoracic Society
BEN HAMLIN, MPH, National Committee for Quality Assurance
ELIZABETH HERMAN, MD, MPH, CDC *
LARRY KLEINMAN, MD, The Mount Sinai Hospital
JASMINE LARSON, MBA, CPHQ, Minnesota Community Measurement *
COLLEEN M. McKIERNAN, MSPH, The Lewin Group
BELA PATEL, MD, American Thoracic Society *
LINDSEY ROTH, MPP, National Committee for Quality Assurance
JOHN SCHALL, MPP, PCORI *

REBECCA SWAIN-ENG, MS, CAE, American Academy of Asthma, Allergy, and Immunology

*** present by teleconference**

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

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

9:07 a.m.

MS. GORHAM: All right. Good morning, and welcome to the Pulmonary and Critical Care CDP Standing Committee Meeting. Thank you for joining us for our PCC measure evaluation and being Members of the Standing Committee.

My name is Shaconna Gorham, and I am a senior project manager here at NQF. We are going to kick off introductions, and later, we will introduce our Co-Chairs, David Lang and Dale Bratzler, but before then, we want the NQF staff to go ahead and introduce themself.

Poonam?

MS. BAL: Hi, I am Poonam Bal, the project manager on the Pulmonary Project.

MS. AMIRAULT: Hi, I am Janine Amirault. I am the project analyst.

MS. GORHAM: Robyn? Yes.

DR. NISHIMI: Robyn Nishimi, I am a consultant here. I was the founding Chief Operating Officer. And I also wanted to make a

note that I recuse myself from 1893. I work closely with a member of that measure's steering committee, so even though I wasn't part of that project with him, I just decided to recuse myself from that one.

MS. GORHAM: All right. Elisa?

MS. MUNTHALI: Good morning. My name is Elisa Munthali. I am Vice President for Quality Measurement at NQF. I wanted to welcome everyone and thank you for being on the Committee.

MS. GORHAM: Helen?

DR. BURSTIN: Great, hi, I am Helen Burstin, the Chief Scientific Officer here at NQF. Delighted to have you here with us today. Ann Hammersmith will do disclosures with you in a moment.

But I'll be popping in and out of the meeting at any point during the day today if you have any questions that Elisa or these ridiculously smart people lined up here can't answer about our process or our criteria. I am

happy to come back up and assist you, but you are in excellent hands, and just want to say thank you.

We have really found that moving to standing committees has been very helpful. I know this is the first rodeo for you guys, but the idea that you'll stay together over time and really kind of oversee this portfolio we have found to be really, really helpful in that ongoing continuity.

So thank you so so much, and with that, I am going to turn it over to Ann to do introductions. Welcome, Ann.

MS. HAMMERSMITH: Thank you, Helen.

I am Ann Hammersmith. I am NQF's General Counsel, and as Helen said, I am here to lead you through the disclosures of interest.

If you recall, when you applied for the Committee, you filled out a somewhat lengthy form, we wish we could make it shorter but we haven't been able to, asking you about your professional activities. And the information you

gave on that form was an ingredient in deciding who would sit on the Committee and who would not. In the spirit of openness, at the first public meeting of a committee, we ask you to do oral disclosures.

When you do this, we don't want you to summarize your resume, please. We just want you to disclose anything that you believe is relevant to your service on the Committee. Just because you disclose does not mean you have a conflict. I want to emphasize that.

So what we're looking for you to do is to disclose any activities that you believe are relevant to the work before the Committee, such as consulting, speaking engagements, grants, research, and so on.

I want to remind you that you sit on the Committee as an individual. You do not represent your employer. You do not represent an organization that you belong to. You do not represent anyone or any organization who may have nominated you to serve on this Committee.

And the last thing I want to remind you of is that people will often say I don't have any financial disclosures, which is, you know, that's fine, great if you don't have financial conflicts, but the way NQF's disclosure of interest process works is that it's -- the disclosures aren't strictly driven by money exchanging hands.

So in other words, you may have served on a committee for your professional society that is relevant to the work before the Committee. You're not paid, but we would look for you to disclose that, but only things that are relevant to service before the Committee.

So what we'll do is we will go around the table, ask you to give your name, who you're with, if you have anything you would like to disclose. If you're on the phone, I will call on you.

So let's start with the Chairs.

CO-CHAIR BRATZLER: Hi. I am Dale Bratzler. I am with the University of Oklahoma,

both the College of Medicine and College of Public Health.

I actually have no formal disclosures at this point, financial or otherwise. I did recuse myself from one of the measures, CMS 0468, which is a measure that I had worked on extensively in the past.

CO-CHAIR LANG: Good morning. I am David Lang, and I have engaged in clinical research with, have received honoraria from, and/or have served as a consultant for Merck, GlaxoSmithKline, Genentech, Novartis, Meda Pharmaceuticals. I am the Department Chair of Allergy and Clinical Immunology at the Cleveland Clinic and Co-Director, Asthma Center. I am involved with quality measures at the Cleveland Clinic, and I also serve as the chair of the Measure Stewardship Committee of the American Academy of Allergy, Asthma & Immunology, and I am one of the two co-chairs of our Practice Parameters Task Force for Allergy Immunology. That is our guidelines development group.

Thank you.

MR. BENSON: Good morning. My name is Ken Benson. I am a patient advocate representing the COPD Foundation. I have no involvement with any of the measures before us other than being a patient.

MS. FRECHETTE: I am not on the Standing Committee. I am Sue Frechette, measure developer representing ATS.

DR. DORMAN: Hi, sorry. Todd Dorman, I work at Johns Hopkins where I do both critical care and am the Associate Dean for Continuing Medical Education.

I am here -- I don't have any conflicts based upon your definitions.

DR. GLOMB: Good morning. I am Brendle Glomb. I am Chief Medical Officer for Superior HealthPlan, which is the largest subsidiary of the Centene Corporation. We do entitlement insurance, Medicaid, Medicare exchange, et cetera.

And I work hand-in-hand with our Vice

President for Quality, so I am -- I am one of those evil-side people who puts these measures into play in the field with managed care. I am a pediatric pulmonologist and neonatologist. I have nothing to disclose.

MS. WEST: Good morning. My name is Chana West. I work for Booz Allen Hamilton. I have no direct disclosures, although I have worked in the past on measure development contracts, but none in this realm.

DR. OHTAKE: Good morning. My name is Patricia Ohtake. I am at the University of Buffalo in the School of Public Health and Health Professions, and I have no disclosures to report.

DR. ELLIOTT: Hi, I am Kim Elliott. I work for the Arizona Health Care Cost Containment System, better known as AHCCCS, which is the Medicaid program in Arizona. I am the Clinical Quality Management Administrator, so I am responsible for all of the clinical programs for the Medicaid program, and I have nothing to disclose.

MS. HAMMERSMITH: Could you use your neighbor's while we get you another mic?

DR. RILEY: Hi. I am Crystal Riley. I work for Baxter Healthcare Corporation as their Senior Manager of Healthcare Policy and Reimbursement. I have no disclosures.

DR. LAMPONE: Good morning. I am Tom Lampone. I work for Florida Blue. I am also a member of the Pharmacy Measurement Development Committee, and I have no other disclosures.

DR. COLLINS: Hi, good morning. Curtis Collins. I am a clinical pharmacy specialist at St. Joseph Mercy Health System in Ann Arbor, Michigan, which is part of Trinity Health as well. I have no disclosures.

DR. DiGIOVINE: Hi, I am Bruno DiGiovine. I am the Division Head for Pulmonary and Critical Care Medicine at Henry Ford Hospital in Detroit. I am a consultant for the Michigan Hospital Association, do quality improvement work for ICUs throughout the State of Michigan, but otherwise I have no disclosures.

DR. BAULDOFF: Hi, I am Gerene Bauldoff. I am a professor at Ohio State in the College of Nursing, and my only disclosure is that I've served as the developer of other measures in pulmonary rehabilitation.

DR. JIMENEZ: Hi, good morning. I am Edgar Jimenez. I am Vice President of Critical Care for the Baylor Scott & White System in Texas. I have served as a board member from the American College of Critical Care, and I was Vice President of the World Federation -- and I still am Vice President of the World Federation of Critical Care.

I currently sit on the ABIM Committee for Critical Care, and I have no disclosures.

DR. GROSSBART: Hi, I am Stephen Grossbart. I am the Chief Analytics Officer at Mercy Health in Cincinnati, Ohio. I have no disclosures. I have served on the Agency for Healthcare Research and Quality, CAUTI technical expert panel, and done some -- received some honorariums from the American Hospital

Association not related to measurement development, and formerly worked with Premier, or did work on a measure development. That was over 12 years ago.

DR. O'BRIEN: Good morning. I am Jim O'Brien. I am the Vice President of Quality and Patient Safety at OhioHealth Riverside Methodist Hospital, so my day job involves reporting on these measures.

I am also a consultant to the Ohio Hospital Association regarding a statewide effort towards sepsis care, and I am the chairman of the board of directors of Sepsis Alliance, a not-for-profit dedicated toward sepsis awareness.

As a result of that, I do a number of speaking and advisory board roles and donate the honoraria back to that charity. That includes engagements with Tenex, Abbott, Medical Simulation Corporation, GE, and Ortho Clinical Diagnostics.

DR. MURRAY: Good morning. Richard Murray, Deputy Chief Medical Officer at Merck.

My disclosures are I'm a full-time employee of a manufacturer. We do manufacture several drugs that are used in testing.

DR. SCHINDLER: Good morning. My name is Christine Schindler. I am a nurse practitioner with a joint appointment at American University and the Medical College of Wisconsin.

I -- the only -- I don't have any financial disclosures. I was an expert consultant for the AHRQ pressure ulcer quality work. That's it.

MS. HAMMERSMITH: Okay, thank you. Now I'm going to turn to the people who are on the phone. I will call you by name.

Susan Pollart?

DR. POLLART: Hi. I am Susan Pollart. I am a family physician at the University of Virginia and the Senior Associate Dean for Faculty Affairs and Faculty Development. My only disclosure is service on the TEVA Respiratory Healthcare Advisory Board that occurred more than four years ago.

MS. HAMMERSMITH: Okay. Thank you. Mitch Harris?

DR. HARRIS: Hi, I am Mitch Harris. I am a health services researcher who works for the Children's Hospital Association focusing mostly on quality advocacy.

I -- disclosure, I have been on a couple of AHRQ time-limited work groups that have focused on the PDIs and some of the other measures that might be up for this group, and also by association the part owner of a company that has -- has developed some of the measures that are under review today, and so I've been recused from a number of the measures for discussion.

MS. HAMMERSMITH: Okay. Thank you. David Stockwell?

DR. STOCKWELL: Good morning. I am a pediatric intensivist at Children's National there in Washington, D.C. Sorry I can't be there today.

I -- by way of disclosures, I also am

partially employed by Pascal Metrics, which is a patient safety organization, but I don't believe there's any conflicts. Thank you.

MS. HAMMERSMITH: Okay. Thank you.

Is there anyone else on the phone that I missed?

(No audible response.)

MS. HAMMERSMITH: Okay. Thank you for those disclosures. Based on the disclosures, do you have any questions of me, or do you have anything that you want to raise with your fellow Committee Members?

(No audible response.)

MS. HAMMERSMITH: Okay. Before I leave you today, just one more reminder, then I promise I'll go away.

In order to make the conflict of interest process work, we rely on all of you. Everybody has a role to play in this. So if during the meeting you think you have a conflict, or if you think one of your Committee Members has a conflict, or if you think someone is behaving

in a very biased way, we're looking for you to speak up. Please don't sit there in silence and then a month later say you know, I think I had a conflict, or I think somebody else had a conflict.

You are always welcome to bring up an issue in open session. If you don't want to do that, you can approach your Chairs, and they will bring it to staff, and it will be taken care of, figured out, or you can go directly to staff.

So thank you, and have a good meeting.

(Pause.)

MS. GORHAM: All right, so we will briefly do an overview of the Pulmonary and Critical Care, you may hear me say PCC just to shorten it, Portfolio. Next slide.

So these -- the next couple of slides are not foreign. We also reviewed them during the orientation call. So today we've convened to evaluate the measures related to the PCC conditions measures that are used for accountability and public reporting for all

populations in all care settings.

The next few slides, you will see the measures that are included in the -- in the portfolio. Those measures in purple with the asterisk next to them indicate the maintenance measures that we will review. There are 18 maintenance measures in this project, and there are four new measures in the project.

The measures that are in orange are part of the portfolio, but they will not be reviewed in this project.

So this first slide indicates that there are four asthma measures and one asthma and COPD measure.

We also have four COPD measures, three pneumonia, and one imaging measure. We have five critical care measures, and then again, at the bottom, you will see the four new measures that we will review.

There are a mix of outcome and process measures in this portfolio, and of course, it is always good to see outcome measures in the

portfolio as NQF puts a premium on outcome measures.

As Standing Committee Members, we ask you to take ownership of the portfolio, so not only the review today, but also because you are experts in the field, as we go through the day, you will start to notice gaps in the portfolio, so we ask you to identify those gaps as well as help us to identify new measures that can make the portfolio more robust.

So as a matter of full disclosure, there are some measures that are part of the portfolio, but they are being retired by the developer. So those are the measures on the screen in this slide as well as the next slide.

Okay. Okay. So our activities and timelines. Of course, over the next two days, we will evaluate the measures and make recommendations for those measures. At the post-meeting webinar, if we are not -- if we do not have the time or for some reason cannot evaluate all of the measures today, we will have a post-

meeting webinar. If we are successful and we finish all of our measure review today and tomorrow, we will not have the post-meeting webinar, but it is scheduled for March 22 at 12 o'clock.

And then we also, after the webinar if we have one, after you all make your evaluation and you have your initial review of the measures today and tomorrow, we will develop the draft report, we being the staff, we'll have a draft report of the evaluation and the recommendations.

That draft report will be posted on April 21. It is a 30-day comment period, so that would end on May 20. And then we will convene you all back together via telephone call where you can review the comments made to the report, and that call will be on June 13. That is at 1 o'clock. And then the draft report will be posted for NQF Member vote.

After NQF Member vote, CSAC will review the information that we -- that you all -- the evaluation from -- from today and tomorrow as

well as the comments as well as the Member vote, and CSAC will convene and review and improve the measures in the evaluation on July 13 and 14. Then the endorsement by the Board would be August 3, and then if we receive an appeal to one of the endorsement decisions, then that will happen on August the 5th to September 30. It is a 30-day appeals process, so an appellant can make an appeal, they have 30 days to do so.

Okay. Okay. So the ground rules for today's meeting. I am not going to read them verbatim. However, as a result of the work group calls that we've had, you all have had a chance to actually look at the measures, so we have definitely prepped for today's call, and the measure discussion script has been provided. It was emailed to you all. But also there was a handout at your seat, and so you also have that in front of you.

So if you all could just take a minute to review the ground rules, and if there are no questions -- just a bit of housekeeping. Only

three mics can be on at one time, so after you finish speaking, if you can cut your mic off. And remember to speak directly in the mic. Our transcriptionist will be taking notes throughout the meeting.

Okay. Okay. So NQF continually strives to improve our Committee meetings, and so measure developers have been invited to our meeting. At the beginning of each measure discussion, they will have two to three minutes to briefly introduce themselves. They will sit to the left of me.

Committee Members, of course, we have assigned discussants for each measure, and so we have lead discussants. You all will begin to explain the measure, start with the criteria and go through that, and then the measure developers will have an opportunity to answer any questions that you all may have or to clarify any information. They will be invited to respond to the questions at the discretion of the Committee, so the conversation is mainly for Committee

Members, but the developers are definitely here to answer any of your questions or to clarify anything.

Committee Members often offer suggestions for improvement to the measures. Suggestions can be considered by the developers for future improvement, but we do expect you to evaluate and make recommendations on the measures based on the information that you have in front of you today.

Committee Members act as proxies for NQF membership, and so we really do value the multi-stakeholder perspective that you all bring to the table.

Okay. So voting for endorsement criteria. Again, we went through these criteria on the work group calls as well as the orientation call, but just to remind you, importance to measure and report and scientific acceptability are must pass. Under importance to measure and report, we have evidence as well as gap, and that is also where we will discuss

disparities in information if provided by the developers.

Reliability and validity of the measure is really the crux of the measure. We want to make sure that the measure is reliable and valid.

Feasibility is also critical. We want to make sure that the measure can be implemented with less burden as possible. We want to look at usability and use as well as we'll look at related and competing later on. We don't have any competing measures in this portfolio, but we do have quite a few related measures.

Okay. I will hand it over to Poonam to discuss the next few slides.

MS. BAL: Okay, so this is where we're getting a little active, voting during today's meeting, so I am sure you all have been waiting for this.

So it is a -- all in-person voting members have a clicker, so please make sure you have one of these blue clickers. The person who

are on the phone will be chatting us their votes, and staff members have clickers designated for them to vote on their behalf.

All voting Members -- sorry, never mind.

Instructions: so for that, please click your -- point your clicker towards Janine over here. I know you're going to have an urge to point it over there because that's the screen that you're going to see, but we need it to go to the system.

When voting, you'll -- your remote will tell you right here in this little screen what your vote was, so if it didn't match up to what you wanted to vote on or you get any sort of error message, please let us know. We're going to do a test run in just a couple seconds, and so we want to make sure everybody's remote is working.

And then also, let's say within the time that you have, you decide oh, never mind, I no longer like my response. You can just click

the number, and whatever your last response was is the one that will be registered, so don't feel that you're stuck with whatever you respond with.

Okay. So with that said, we're actually going to do a test run. Janine, could you just pull up an evidence slide, please?

(Pause.)

MS. BAL: So I ask that everyone that is on the phone please message us either 1 or 2 on the chat feature, and everybody in the room please also click either 1 or 2 on your remote, and if you don't see the number, please let us know.

(Pause.)

MS. BAL: So we did not receive one on the chat. Can everybody in the room please just click their number again just so we can make sure it works.

MS. GORHAM: David, we're still waiting for your vote through chat.

DR. KAZEROONI: This is Ella Kazerooni. I am trying to chat, but when I --

the send button does not seem to be sending.

MS. BAL: Actually, Ella, we received your vote. Thank you.

DR. KAZEROONI: Oh, you did? Thank you.

MS. BAL: Yes.

DR. STOCKWELL: Hey, I am sorry, it's David. But I've sent it a couple times. You haven't received it?

MS. BAL: Not yet. Could you just -- Operator, could you work with David to make sure he's able to chat us his responses?

(No audible response.)

MS. BAL: All right. So Janine, could you do a test run of -- showing the votes?

MS. AMIRAULT: Okay, so as you can see, we have twelve -- twelve votes for yes, and nine for no, with a total of 21.

MS. BAL: So that went pretty efficiently. No one had any error message. We'll work out what's going on with David on line, and if not, we'll get another method for

getting him to vote.

Okay. So we'll move forward to the next slide, Janine. Next slide, please. Over here.

Thank you. Okay. So achieving consensus: so a pass or a recommended vote will be anything greater than 60 percent, and that does not include 60 percent, only greater, and a yes vote would be considered the sum of a high and moderate if we have a for-decision point. If we just have a yes or no question, then it would just be yes.

Consensus is considered not reached if we are between the 40 to 60 threshold, and again, that includes 40 and 60 in that number. If we do not get at least 40 percent, then it's considered a do not pass and not recommended. Again, these numbers really matter for the first two criterion because those are must pass.

For the second two, if -- if they do not pass, we do still move forward, but that's included in the draft report, and we will ask

public comment on that.

And quorum is considered 66 percent of the Committee. We are definitely way beyond that number right now, and we'll continue to monitor as that -- as we go through the meeting. Next slide, please.

Okay. So the last little aspect is related and competing. During the work group calls, we did have some Committee Members that wanted to review a measure based on another measure that was very similar. However, we really like to review the measure based on its merits solely.

So as the measure is with basically just your knowledge of the -- the community, but not the knowledge of other measures, how does it pan out? And then once, if it does reach endorsement, we will compare that measure to the -- any other similar measures that are also recommended for endorsement. We want to make sure that they are -- none of the other thoughts are conflicted with that, and that point, as

Shaconna mentioned, we do not have any competing measures, so you will not be making a best-in-class decision.

However, you will be making recommendations on harmonization. You will be making the decision on if developers should harmonize further or if you feel that the way that the measures are, even though that they're related, they have their own unique point of view and they do not need to be harmonized any further.

All right. Were there any questions about that?

(No audible response.)

MS. BAL: Okay, perfect. Next slide, please.

So with that, I'll give it to the Co-Chairs. Shaconna, did you want to say something?

MS. GORHAM: Just real quickly, if -- to the extent possible, I know it won't happen all the time, but to the extent possible, if you could remain in the room until break, because we

want to make sure we maintain quorum so that when we vote we have the needed number in the room, and so quorum for this Committee is 15, so to the extent possible, I know we all have to step out when we need to, but to the extent possible, thank you.

And we're going to turn it over to David for the first measure.

CO-CHAIR LANG: Thank you, Shaconna.

Prior to getting into the first measure, Ella has joined us on the phone. Ella, could you please elaborate for us on your disclosures?

DR. KAZEROONI: I have no disclosures.

CO-CHAIR LANG: Thank you.

So with that, we will proceed to consider the first measure, which is 2852: Optimal Asthma Control, Minnesota Community Measurement. Do we have Jasmine on the phone?

MS. LARSON: Yes, I am here.

CO-CHAIR LANG: Please, Jasmine, would you like to give us a brief two- to three-minute

summary of the measure?

MS. LARSON: Sure, I'd be happy to. Thank you.

So this measure is a patient-level all-or-none composite. As you likely already know, a version of it was brought to this Committee in 2012. That was a three-component all-or-none composite measure. However, due to a lack of a strong and consistent body of evidence for the third component, which was a written asthma action plan, the measure did not pass the evidence criteria at that time.

In follow-up, we reconvened our measure development work group and ultimately ended up removing that third component from the measure, so we're presenting to you today a two-component composite, one of which utilizes patient-reported outcomes based on the use of any of the three tools, four tools actually, specified, which all have been validated with -- to have clear-cut points that define control.

And the second component is a patient

recall of the last 12 months of emergency department or hospitalizations due to asthma, and for numerator compliance for that component to meet the target, there has to be zero or one events of that nature.

This measure is reported -- attributed to the provider, reported at the clinical -- clinic site or large-practice level.

And that's all I have for you this day. Thank you.

CO-CHAIR LANG: Thank you, Jasmine.

We have two discussants, Steve Grossbart, Christine Schindler. Please take it away.

DR. GROSSBART: We didn't rehearse, so we don't know who is going to go first.

Well, I will introduce the measure then. As already noted by the developer, this is an all-or-nothing composite consisting of two components.

The -- the -- this is a new measure. Again, it was considered in an earlier version

before the revisions by this -- the predecessor committee of this Committee. It is an outcome measure, and I just, I am trying to go through my guidelines.

So I think we want to move to the discussion of the first components of the -- the measure evaluation.

DR. SCHINDLER: Right. So I think that --

CO-CHAIR LANG: So initially, we should be discussing the evidence --

DR. SCHINDLER: Right --

CO-CHAIR LANG: -- and then --

DR. GROSSBART: Which is the evidence --

CO-CHAIR LANG: -- then --

DR. GROSSBART: -- that's --

CO-CHAIR LANG: -- yes --

DR. GROSSBART: -- what I meant, the first component of the evaluation would be the evidence, correct?

CO-CHAIR LANG: Yes.

DR. GROSSBART: Okay.

DR. SCHINDLER: So we appreciated, it sounds like the developers did take into consideration the comments from the last group, including there wasn't clear evidence that if they had fewer ED visits and better report of control, why we also needed to have the written asthma action plan in place. There weren't good data to support that.

One of the questions we did have this time around was that there really wasn't new evidence related to the two composite -- or the two measures that were left. There's some graded clinical evidence for component one, but there was no clear graded evidence for component two, which was the self-report of ED visits.

DR. GROSSBART: Or hospitalizations.

DR. SCHINDLER: Or hospitalizations, right, thank you.

DR. GROSSBART: One of the -- the points of discussion within the work group was that the components have a -- a -- draw heavily

on the clinical guidelines, and there is evidence for them. These are -- these are the -- the first numerator is -- is asthma well-controlled as defined by the most recent asthma control tools, so there's evidence for that measure, and then the second one is -- is the occurrence of a -- more than one ED or inpatient admission visit, inpatient visit, in the course of the measurement period.

What their measure developer has not provided is the evidence base for the overall composite.

CO-CHAIR LANG: Closer.

DR. GROSSBART: Talk louder?

CO-CHAIR LANG: Closer to the mic, yes.

DR. GROSSBART: Okay, closer to the mic, sorry.

So we don't -- we don't -- the measure itself hasn't been tested or hasn't -- there hasn't been the evidence base as robust in that area. That was one of the points raised in our -

- our initial evaluation.

And -- and actually, the other point was the ED and inpatient visit intuitively makes sense, but there was no graded evidence for that component. I misspoke earlier. So there's no graded evidence for ED/inpatient visit as being evidence of -- of well-controlled asthma, and there was graded evidence for component one, which is the Asthma Control Plan Questionnaire.

CO-CHAIR LANG: So before reflecting this back to the group, just a -- a mention that if you wish to comment, those around the table, oops, please put your name badge in this -- or, you know, reorient yourself 90 degrees, and that way, I will know to call on you.

For those of you on the phone, there is a mechanism whereby you can quote unquote "raise your hand" so that we know that you have something to say and contribute.

So with that said, are there any other comments from the group? Please.

DR. DiGIOVINE: Yes, I had two

questions that I couldn't quite figure out.

One was in terms of the Asthma Control Test, I'm not sure I get where the -- where that measure would be gotten from and how it would be scored if there was not one that was available in the last I guess 12 months.

The second was do -- do the developers provide any sense of whether a self-controlled -- a self-reported ER or hospitalization actually correlates with actual ER and hospitalizations? To what degree is patient recall accurate?

MS. LARSON: This is the developer. Would you like me to respond?

CO-CHAIR LANG: Please.

MS. LARSON: Okay.

Well first, I want to comment on the evidence in that this is an outcome measure, and the evidence requirement for that is just that a rationale supports the relationship of the outcome to a process or structure of care, so graded evidence is not part of the requirement as we understand it for endorsement.

In regards to the Asthma Control Test, we are looking for a tool result in the 12-month period, and the absence of a tool result is judged as not in control. It is -- it is a numerator miss, as these patients to come into the denominator do require a face-to-face visit with the provider, and if they are having a face-to-face visit and they have a diagnosis of asthma, the expectation is that they are assessed for control.

In regards to the evidence for patient recall, there is strong evidence to support that patient recall is accurate in the last 12 months regarding emergency room and inpatient hospitalizations.

CO-CHAIR LANG: Thank you for your response.

MS. LARSON: You're welcome.

DR. GLOMB: With regard to the second numerator and the patient recall -- with regard to the second numerator and the patient recall of their emergency room visits, we routinely contact

a patient within 72 hours of receiving a claim for an emergency room visit with asthma, and that claim of course can be anywhere from minutes to -- to three months later.

Nevertheless, the patient recall even on that -- on that call directed at a specific date with -- with my company knowing that is -- is often very difficult to -- for them to -- to recall. So again, I am concerned that without some sort of a verification, i.e. a claims-based database for the emergency room visit, that this -- this measure is subject to a lot of -- of vagary.

CO-CHAIR LANG: All right. I am actually going to call on myself now.

And I have a question for the developer, and it is this: you have, among the asthma control instruments in the numerator, you have the Asthma Control Test, including the Childhood Asthma Control Test; the Asthma Therapy Assessment Questionnaire; and the Asthma Control Questionnaire.

The ACT and the ATAQ have been well validated, and they are four-week instruments. The Asthma Control Questionnaire I have seen used largely in clinical research studies, and the recall period is only one week as I recall based on that instrument, so I am wondering whether you can help us reconcile this in terms of the use of the Asthma Control Questionnaire.

MS. LARSON: The Asthma Control Questionnaire is validated with a control cut point defined as 0.75. It is true that there is an indeterminate range of 0.75 to I believe 1.25 or 1.5 for determining truly out of control or in control, but the group felt strongly about using the 0.75 cut point, as that is what is specified in the NHLBI Guidelines where they talk about the identification -- let's see here -- classifications of asthma control.

The three tools selected were based on those guidelines. Those are the three and only three that are identified, and those are the cut points identified as well.

CO-CHAIR LANG: Thank you. Are there any other comments? Please.

DR. GROSSBART: So just one real quick.

Can you tell me exactly how this composite is calculated? I mean, so you have an Asthma Control Questionnaire, and you have this recall of ED or inpatient visit, so how is the composite actually calculated? I see the rates that you report in the metric, but how is it calculated?

MS. LARSON: Sure, so it's an all-or-none composite at the patient level, meaning that all patients who have that -- established patients who have that face-to-face contact with an eligible provider and the diagnosis are in the denominator, and then to be in the numerator, it's an "and" statement. The patient needs to have an Asthma Control Tool result in control and report less than two emergency department visits and/or hospitalizations due to asthma in the last 12 months.

Does that answer your question?

DR. GROSSBART: Yes. I -- I've just worked with all-or-none composites in the past. I think this one is okay, but sometimes if you have multiple components of a composite, particularly the all-or-none, if the numerator of one of the two components or multiple components is more than others, it can skew the results. But this one, I don't think that's a problem. I just wanted to make sure I understood how it was calculated.

MS. LARSON: Okay. Great.

DR. GROSSBART: I do have a point of -- just a question for the staff.

So as an outcome measure, the evidence bar is different than from a process measure. But this is a composite that includes a process measure plus an outcome measure, and what is -- what is the -- what is the bar for the evidence for the process measure that is -- is a subset of the overall measure?

And then the other comment is that in

-- in the past, I don't know that this is absolutely true, but in the past, the composites I've seen that have been approved, including one developed by this -- this -- the -- this developer, consisted of all NQF-endorsed measures, and neither of these are NQF-endorsed, and should that be a factor in our evaluation or not? Should a composite consist of sub-components that have been endorsed by the NQF?

MS. BAL: So we've put up a slide for you for the voting so you'll understand that the vote is different based off of the fact that it's a composite. You'll be looking for slightly different things.

So unfortunately I can't read from here, but basically, you are seeing how the composite works together and if there's evidence to support it, and I'll see if anybody from NQF staff want to add any more detail.

MS. GORHAM: Robyn? Could you answer his question regarding the difference between the process and outcome?

DR. NISHIMI: So the process measure has to stand by itself, and the outcome measure has to stand by itself. As to whether --

PARTICIPANT: And this is the tool developer. I -- I am a little confused about the classification of one of these components as a process measure.

DR. NISHIMI: I was just --

CO-CHAIR LANG: I would agree.

DR. NISHIMI: -- I was just going to say --

CO-CHAIR LANG: It's an outcomes measure.

DR. NISHIMI: -- they -- they call both of them outcome measures, so in that case, each component as an outcome measure --

CO-CHAIR LANG: So asthma --

DR. NISHIMI: -- needs to stand --

CO-CHAIR LANG: -- control, she being asthma control is an outcome.

DR. NISHIMI: -- as itself. So they characterized it --

CO-CHAIR LANG: And I don't think of it as --

DR. NISHIMI: -- as an outcome measure.

CO-CHAIR LANG: -- giving the questionnaire. Got it. I stand corrected.

DR. NISHIMI: Oh sorry, and Stephen, to your -- your other question about whether each had to be individually NQF-endorsed, no, they don't, but they do have to stand alone on themselves, and you will consider them in that manner.

Were there any other questions on how you're going to consider the evidence here?

(No audible response.)

CO-CHAIR LANG: Are we ready to vote?

MS. BAL: Okay. So as a reminder, everyone that's on the phone, please chat your response. We're now voting for the evidence, the composite for 2852. Your options are 1 high, 2 moderate, 3 low, 4 insufficient, and we're looking at this composite, specifically

articulates and is logical, and that includes all components and rationale.

DR. STOCKWELL: Could you say the choices again please?

DR. NISHIMI: Repeat the choices.

CO-CHAIR LANG: The choices are 1 high, 2 moderate, 3 low, 4 insufficient.

(Pause.)

MS. BAL: Sorry for the delay. We're still trying to get the logistics of having people on the phone. Give us one second.

(Pause.)

DR. NISHIMI: She has to reset the system, basically.

(Whereupon, the meeting went off the record at 9:57 a.m. and resumed at 10:02 a.m.)

DR. NISHIMI: Okay. We're ready to get started again. Sorry about that. We're going to revote on that last slide. So, as soon as we get everyone settled, we'll revote.

For folks who are on the phone, we have your votes in chat, so you don't need to

retransmit.

(Pause.)

MS. BAL: All right, everyone. Thank you for your patience while we had some technical difficulties. So, we're actually going to start the vote on evidence health outcome. And the composite vote will come after the Outcome and gap vote. Is that clear to everyone? So, we will need to have -- they are different selections, so everyone will need to revote.

So, now we are voting on evidence for 2852, health outcome. Your options are 1 yes, 2 no. People on the phone, we do need you to revote. Again, the options are 1, Yes, 2, No.

And so, the vote is that the rationale supports the relationship of the health outcome to at least one healthcare structure, process, intervention or service. And the responses are 1 yes, 2 no.

DR. DiGIOVINE: Can I ask a point of order? I'm sorry. Since it's a composite measure, I know this was probably already asked,

but I'm still -- so if you think for one of the composite answers the answer is yes and for one it's no, how would you answer this question? Because it says at least, so I'm not sure I understand how I would answer this in a composite measure. I guess no one else does.

CO-CHAIR BRATZLER: I mean, I'm reading this as the measure is Optimal Asthma Control. That is the outcome that is be measured. There are two components to the measurement, but the measure is Optimal Asthma Control. So, is that linked to at least one healthcare structure, process, intervention or service?

DR. DiGIOVINE: Thank you.

DR. NISHIMI: And you will have a chance to vote on the composite as a whole.

MS. BAL: Okay. Does everybody feel comfortable voting now? Okay. So please put in your votes. Again, the options are 1 yes, 2 no, for evidence for 2852. Yes, please point at me, not Janine this time, with the remote.

Okay. So, the results for 2852 for

evidence health outcome is 100 percent yes, zero percent no. And we can move forward to the discussion on gap.

DR. SCHINDLER: On this measure, I think the developer did a nice job describing the gap. They're certainly -- they've been using this composite measure in Minnesota and they identified -- it was optimal asthma control is 47 percent in adults and 56 in children. They also -- there appears to be some racial, language, and ethnic gaps that they were able to identify. So I felt like they did a nice job on outlining the gaps.

CO-CHAIR LANG: Is there any additional discussion?

MS. BAL: Okay. So we can vote on gap. People on the phone, please put your vote in. So, performance gap for 2852; the options are 1 high, 2 moderate, 3 low, 4 insufficient. Again, the options are 1 high, 2 moderate, 3 low, 4 insufficient. And you're voting on performance gap, the data demonstrates considerable variation

or overall less than optimal performance across providers and/or population groups, and this includes disparities in care. Okay.

(Voting.)

MS. BAL: So, the results for performance gap for 2852 is 73 percent high, 27 percent moderate, zero percent low, zero percent insufficient. And we can move forward to the composite vote.

So, again, this is the composite vote. This is the whole -- the composite as a whole and your options are 1 high, 2 moderate, 3 low, 4 insufficient. And just so everybody in the room knows, the magical number we're looking for is 22.

DR. NISHIMI: Ella, we're still waiting for yours in the chat box.

(Voting.)

MS. BAL: Okay. Thank you for your patience. So the results are 14 percent high, 45 percent moderate, 36 percent low, five percent insufficient. And we can move forward.

Actually, no, I'm sorry, we're consensus not reached, my math is a little slow. Fourteen and 45, yes, so we are in consensus not reached, however, that just means that we will ask for comments from the general public, but we do continue to move on with considering this measure and voting on the next sections. Thank you.

CO-CHAIR LANG: Reliability?

DR. GROSSBART: So, in the area of specification and reliability, there were concerns about the specifications regarding the second component; this is patient recall of one or more ED or hospital visits in the course of a 12 month period. And, the reliability testing of the measure we did feel demonstrated sufficient validity so that you can make conclusions about quality, but we were a little -- there was -- the main validity of the developer's testing of this measure was to see if there was a high correlation between clinics that performed well on the developer's optimal diabetes care measure and if that correlated with performance on the

optimal asthma care measure.

Correlation, we did not believe was very strong, 0.62 for adults and 0.66 for children, we felt those were on the weak side; and those were major points that I wanted to raise. Just scrolling through the -- in terms of the measure, other aspects of the validity testing. Missing data is not a challenge.

Exclusions of the measure, which include -- the exclusions are appropriate, such as patients with COPD. There is a risk adjustment within the measure, which adjusts for depression, tobacco use, and tobacco exposure. The risk model didn't seem to be that strong in explaining away variation. Those are the --

CO-CHAIR LANG: Risk adjustment would be in the area of validity.

DR. GROSSBART: Am I jumping ahead here?

CO-CHAIR LANG: Yes.

DR. GROSSBART: I'm sorry.

CO-CHAIR LANG: That's okay, Steve; go

ahead.

DR. GROSSBART: Okay. Well, then I'll just turn it over to my partner here --

DR. SCHINDLER: Right.

DR. GROSSBART: -- for additional comments.

DR. SCHINDLER: I think you hit all the main reliability points that we had discussed. I had a couple for validity, but --

CO-CHAIR LANG: Are there any other comments from members of the committee concerning reliability?

DR. DiGIOVINE: Yes. We brought it up earlier, I just -- in terms of validity, it would seem like this would be the --

CO-CHAIR LANG: We're not there yet.

DR. DiGIOVINE: Okay, then I will hold my point.

CO-CHAIR LANG: Dale?

CO-CHAIR BRATZLER: So just one real quick question and perhaps the developer -- I mean, there's been a lot of question raised about

the recall of ED or inpatient admission. So, has the developer considered using -- since I suspect this would be calculated at the plan level often -- claims data or other sources, rather than patient recall for that component of the composite? The ED or inpatient visit?

MS. LARSON: Yes. This is Jasmine. Because our data source is the medical record provided and the data is provided by the provider, the practices themselves, we have not explored merging data from different data sources to get at ED visits or hospitalizations due to asthma. Also, and part of that was due to our evidence review and the articles we found that demonstrated that recall was a reliable way to capture these events.

CO-CHAIR LANG: Thank you, Jasmine. Are there any other comments, questions concerning reliability? Please, go ahead Bill.

DR. GLOMB: Yes. Just to follow up on what Dale said, again, I commented on this earlier, I don't know at what point we can

interject suggestions, but I think it's time to interject that that data sources needs to be different than the provider record. All of us as providers know that even when we ask, we don't always incorporate the answer or get an answer in the medical record as to have you been in the emergency room since our last visit?

And so I'm just very concerned that that then would be used as a marker of how well the primary care physician, for instance, might be taking care of the patient's asthma, when in reality they simply may not have that number or the number might be incorrect based on patient not -- misremembering, shall we say.

CO-CHAIR LANG: The other issue, just to mention this, I recall when this quality measure was addressed a few years ago that we discussed the issue of hospitalization versus an emergency department visit not being apples and apples in terms of what it means with regards to the patient and the risk domain. That there's a subjective component in terms of whether one goes

to the emergency department, but then there are more objective criteria as to whether someone gets hospitalized.

So, Jasmine, I'm wondering whether you can comment on that issue, which was raised previously and I'm raising it again at this time.

MS. LARSON: Certainly, I'll do my best. Again, the measure development work group relied heavily on the NHLBI Guideline. And in their classification of asthma control, when they are evaluating potential risk of exacerbations, zero to one events, whether they be ED or hospitalizations, were defined as well controlled in the last 12 months.

MS. BAL: Okay. So, we're now going to vote on reliability for 2852. The options are 1 high, 2 moderate, 3 low -- I'm sorry.

CO-CHAIR LANG: Excuse me.

MS. BAL: Sorry, hold on one second on the vote.

CO-CHAIR LANG: Excuse me, we have one more comment.

DR. MURRAY: Well, I'd say it's a question. Are there any NQF endorsed measures that are currently in use that have been based on this type of a recall construct? And, if so, are they working well for improving quality? Do we know anything like that?

DR. NISHIMI: Not in the pulmonary portfolio, but I couldn't speak to the rest of the portfolio frankly.

MS. BAL: Not as far as we know.

DR. MURRAY: And then as a follow-on question, did the developers consider an alternative indicator that perhaps might be easier to remember or might actually be reported in the chart, which would be oral prednisone as an alternative to ED and hospital visits?

MS. LARSON: Actually -- so, this is Jasmine. The measure development work group did consider that and the evidence showed that recall of oral prednisone use was actually poorer than recall of ED and hospitalizations. And I will say, this measure has been in use in Minnesota

for upwards of four or five years and this information is well documented in the medical record and validated through audit.

DR. NISHIMI: And, actually, I do want to speak to the question about -- rolling back my tape. There are obviously a whole body of measures on patient reported outcomes that NQF has looked at and there have been vaccination measures that rely on whether the patient reports they received the vaccination, for example, in the case of a pneumococcal vaccination.

MS. BAL: There was a question about if the measures are working well and Robin responded that they are currently endorsed, but we don't have more information than that. Just for the record. Okay. Were there any other questions?

MS. WEST: Thank you. I guess I'm just a little bit conflicted. So, the numerator indicates that it's looking to estimate the number of emergency room visits. Looking at the code list, without knowing all of the CPT codes, are the inpatient admissions assuming that the

patients are coming through the ED?

MS. LARSON: So this is Jasmine. I believe that question was directed towards me.

MS. WEST: Yes.

MS. LARSON: So, again, both of these components are patient reported, including the ED and inpatient events. In inpatient hospitalization -- I'm sorry, it's an ED visit without an admission and/or an inpatient hospitalization regardless of how they arrived there.

MS. WEST: But then it's going into the numerator estimating the number of emergency room visits.

MS. LARSON: It's not an estimate, it is the patient's report of the number of events in the previous 12 months.

MS. BAL: I think Stephen had his card up.

DR. GROSSBART: In the course of this debate, it struck me that since NQF measures are about being able to publically report performance

by a provider for a level of accountability, is there a potential unintended consequence of providers not asking the question, do you recall any visits, and not entering it? So, therefore, I mean, if you're going to be up on the website in Minnesota -- I mean I know that would be not typical, but it could happen, especially if you've known a patient's been in once, why ask again, and get the public dinging.

Whereas, if we used claims data, it would be -- with an admission or visit for a diagnosis of asthma, it would be much more reliable. And again, you did mention the results in Minnesota are audited, but what is the rate of patients who don't have -- whose number of visits to the ED or inpatient is inconsistent with what the claims record is? Have you tested that?

DR. SCHINDLER: And, Stephen, this I think dovetails with a question I had with the validity, is that -- Jasmine, this kind of goes with Stephen's question, but you had said there wasn't a problem with missing data, but there

weren't -- if you could talk about that a little bit more. It just wasn't further explicated and I think it gets to this exact point around data capture.

MS. LARSON: Sure, I'll do my best. So if the patient's response to the question is not documented in the medical record or if documentation of if the question was asked is not in the medical record, that counts as a numerator miss. So, the patient remains in the denominator, but does not meet the target of the measure. And upon validation audit, the auditor needs to be able to confirm the number that was reported in the data submission against the medical record itself.

So any missing data, including a missing assessment or the question not asked counts as a numerator miss, but as I stated, the patient remains in the denominator. And, I'm sorry, I think you had another question and now I've lost track of it in my mind.

DR. GROSSBART: Correlation between the

self-reported data and the claims.

MS. LARSON: Right. So in Minnesota we don't have an all-payer claims database, so we don't necessarily have a data source to be able to provide all of the complete claims history. And then the matching -- the patient-to-patient level matching of the data would also be problematic. And that's then also compounded by the timing of when that would be available.

Our reporting from the medical groups, from the practices themselves, has a data submission that occurs six to 12 weeks after completion of the measurement period dates of service. And, as I'm sure you are all well aware, claims data would not be ready for that kind of comparison or analysis for quite some time lag after that point.

CO-CHAIR LANG: Chana, Rich, did you have questions or comments? Anyone else?

MS. BAL: Okay. Thank you, everyone. So, now we're voting on reliability for 2852, the options are 1 high, 2 moderate, 3 low, 4

insufficient.

Again; 1 high, 2 moderate, 3 low, 4 insufficient. And we are voting on precise specifications and that the testing is appropriate and has a scope with adequate results.

DR. NISHIMI: David, we're still waiting for your vote in the chat box. David? Okay, we got it.

(Voting.)

MS. BAL: So, the results for reliability for 2852 is zero percent high, 50 percent moderate, 41 percent low, nine percent insufficient. That does put us in the grey zone, and as mentioned earlier, we will continue discussing this measure.

CO-CHAIR LANG: Validity, Christine, Steve?

DR. SCHINDLER: Jasmine, I think we touched on a little bit of the validity in terms of missing data. One of the other questions I had for you was in this section about exclusions.

You had documented that cystic fibrosis, COPD, emphysema, and acute respiratory failure were exclusions and then a couple bullets below -- as well as certain eligible populations that consistently less than one percent of the total population met these criteria, which just -- there was some dissonance for me around that because we know that the incidence of those diseases are probably -- I would anticipate Minnesota is not different than the rest of the country -- if you could just touch on why it was consistently less than one percent that met those exclusion criteria.

MS. LARSON: Sure. This is Jasmine. I apologize for that lack of clarity. So, we do ask providers to share an exclusion file around the allowable exclusions. The required exclusions would include the other respiratory conditions, is not included in that one percent.

DR. SCHINDLER: So, in terms of validity then, how many or what percentage of the patients are excluded?

MS. LARSON: What percentage of patients are excluded due to the respiratory conditions?

DR. SCHINDLER: Yes.

MS. LARSON: We don't have that information.

DR. SCHINDLER: Okay.

MS. LARSON: They're generating a patient list based on query of their system with the inclusion criteria and then excluding these diagnoses up front. But the measure development work group felt pretty strongly about excluding these conditions given that the control tools were not validated on patients with those comorbid conditions and control of asthma symptoms, they felt it was difficult to assess the symptom burden and isolate asthma from these other respiratory conditions.

DR. SCHINDLER: Okay. Thank you.

DR. GROSSBART: At this juncture, we're also evaluating the risk adjustment model and the outcome is adjusted by the following measures:

gender, age, zip code, race, ethnicity, country of origin -- if I'm reading this correctly -- primary language, and insurance product, as well as comorbidities of depression, tobacco use, and tobacco exposure. The risk adjustment model only shows that insurance product, Medicaid, Medicare, commercial status, et cetera, is statistically significant.

And then also, I have a question of why tobacco use is considered a risk adjuster? And this is more of a comment to the developer; tobacco use is part of the optimal diabetes care, are you using tobacco? Why wouldn't tobacco use be part of the optimal asthma care as a part of the composite measure as opposed to a control variable in the risk adjustment model? It just strikes me as incongruous. That's a comment.

CO-CHAIR LANG: Anything else on risk adjustment?

DR. GROSSBART: I think that covers our major comments on this section.

CO-CHAIR LANG: Thank you, Christine,

Steve. Any other comments, questions from other members of the committee? Going once, going twice. Thank you. Poonam?

MS. BAL: Okay. Voting for validity for 2852 is now open. The options are 1 high, 2 moderate, 3 low, 4 insufficient. Again, the options are 1 high, 2 moderate, 3 low, 4 insufficient.

You're voting on the specifications are consistent with evidence, that there's an appropriate method and scope with adequate results for the testing, and that the following threats are addressed: exclusions, risk adjustment, meaningful differences, compatibility with multiple specifications, and missing data.

(Voting.)

MS. BAL: Sorry, so the results are zero high, 14 moderate, eight low, zero insufficient. And percentage-wise, we do hit the requirement for a pass on this requirement. So, now, we'll vote on the composite portion.

DR. KAZEROONI: I'm sorry, can you

repeat the results of the voting?

MS. BAL: One second.

CO-CHAIR BRATZLER: It was 64 percent --

MS. BAL: Yes.

CO-CHAIR BRATZLER: -- so it passed.

MS. BAL: Yes. And so, for numbers wise, 14 moderate, eight low. Okay. So we're now voting on scientific acceptability of the measure properties as a composite. And so the options are 1 high, 2 moderate, 3 low, 4 insufficient. Again, 1 high, 2 moderate, 3 low, 4 insufficient. And again, you don't have to wait for me to finish talking, just a way to give you the information, so we can continue to vote whenever you would like.

And the composite for this, we're looking that empirical analysis support composite construction and that composite measures fit quality, construct, add value, and so on.

(Voting.)

MS. BAL: Okay. So the number results

are zero high, 13 moderate, nine low, zero insufficient for the composite scientific acceptability for 2852. And for percentages, we are in the grey zone for this section, but we continue to move on to feasibility.

CO-CHAIR LANG: Discussion of feasibility; Steve, Christine?

DR. SCHINDLER: Jasmine, it looks like this has been really feasible in Minnesota, you've done a nice job collecting. All the data elements are in the EHR. I guess my question for you was: do you know are these measures, particularly the asthma control measures, being used nationally? If not, how hard would that be to implement? Just, are they the current standard of practice everywhere?

MS. LARSON: Thank you for the question. This measure, the composite measure, has been included in PQRS for this program year. This is the first time that it has been implemented nationally and I don't have any information about the feasibility of it beyond

its current implementation.

CO-CHAIR LANG: Jasmine, this is David Lang. I have a question. In the feasibility section, you state all data elements are in defined fields in electronic health records. Yet on the page 1 of the measure information, you state one of the data sources as being paper medical records. Could you please elaborate on that?

MS. LARSON: Certainly. So, our data collection and submission process in Minnesota is one that is supportive of medical groups at all stages of EHR implementation, including those that are on fully paper charts all the way through those that have a fully implemented EHR with discrete data capture. And so, we have specified it so that manual extraction can occur and those groups can participate in reporting for these measures, but we also know that the majority of practices in the state of Minnesota have done builds in their EHRs to be able to capture the data elements discretely so that they

can programmatically extract the information.

CO-CHAIR LANG: So, a brief follow-up. So then, if your health service utilization component of the measure is by patient self-reports, so then those data presumably are being entered into Smart Fields in the electronic medical record so they're retrievable?

MS. LARSON: That is correct.

CO-CHAIR BRATZLER: Jasmine, what -- tell me about the Minnesota public reporting. I mean, what's the teeth that makes all the practices do this? I can see it being done in a health plan level or other things, but, I mean, do you have every practice in the state? What are the teeth that make people actually do this, to add these data fields to the electronic medical record or report this?

MS. LARSON: Sure. I mean, we are the first to admit that Minnesota is unique and we have had voluntary data reporting since 2004, but we also do have the benefit of, since 2004 actually -- I'm sorry, effective 2008,

implemented in 2009, we had a Minnesota Department of Health state mandate for quality reporting for physician practices.

So, I hesitate to call that teeth because there isn't necessarily penalties or enforcement of that, but it is a state rule as well. But prior to that state rule, we still had 70 percent of medical practices voluntarily submitting data to us.

MS. BAL: Okay. So we're now going to vote on feasibility for 2852. The options are 1 high, 2 moderate, 3 low, 4 insufficient. Again, 1 high, 2 moderate, 3 low, 4 insufficient. And for Feasibility, we're looking at that the data is generated during care, there are electronic sources, and that the data collection can be implemented.

(Voting.)

MS. BAL: Okay. The results for 2852 feasibility are seven high, 12 moderate, three low, four insufficient. And with the percentages, we are actually passing this measure

and we can move forward to usability.

CO-CHAIR LANG: Usability is the extent to which consumers, purchasers, policymakers, use or could use the performance results for accountability and performance improvement activities and would include discussion of unintended consequences. Please; Steve, Christine?

DR. GROSSBART: So, this is currently in use, as noted, in Minnesota. It's available on their website, each practitioner, each group is -- sees their overall score. It has been shown to lead to improvement of care in Minnesota over time. The unintended consequences appear to be minimal.

It does not appear to be burdensome for clinicians to collect in Minnesota. And as noted, this has been incorporated in PQRS, so it's gaining traction as a quality measure nationally, or has the potential to gain traction as a quality measure nationally.

DR. SCHINDLER: I don't have much to

add. The Minnesota group is really using this quite extensively in both accountability, payment, quality measures within the state. They seem to -- yes, do a great job using it and publically reporting it and it's easy to find.

MS. BAL: Okay. So we're now voting on the usability and use of 2852. The options are 1 high, 2 moderate, 3 low, 4 insufficient. And you're looking at if there's accountability and transparency, if there's been improvements seen, and if the benefits outweigh the evidence of unintended negative consequences. Again, the options are 1 high, 2 moderate, 3 low, 4 insufficient.

(Pause.)

MS. BAL: Okay. So we're missing somebody in the room for one more vote. Did someone step out? Okay. You have two seconds. All right. So we're just going to go ahead and move forward.

(Voting.)

MS. BAL: Oh, we got the last vote in.

So, the results for usability for 2852 are nine high, 12 moderate, one low, zero insufficient. And for percentages wise, we can move forward. Okay. So now we're going to vote on the overall suitability for endorsement for the full composite measure. And your options are 1 yes, 2 no. Again -- oh, I'm sorry, there's a question. Hold on one second.

DR. O'BRIEN: Just a quick question. I know this isn't part of the voting, but the comparison of related or competing measures and harmonization, the developers mention that there aren't relations to other NQF measures. I can't help but notice that we have multiple other measures looking at admission rate for asthma, ED visits for asthma, so I'm unsure why these weren't identified as potentially related measures by the developer.

MS. LARSON: This is Jasmine. I believe those are new candidate measures as well, are they not?

DR. O'BRIEN: I don't believe so.

Those are in maintenance.

MS. BAL: So if you look, we have identified some measures as related to the optimal asthma care, we do want to make sure that -- not all measures are considered related. If they don't have the same population or don't have the same focus, they wouldn't be considered related.

The reason that the two measures that were identified as related is because an aspect of this measure is already covered in those measures. So, those are more just kind of similar concepts, but not the same focus.

DR. O'BRIEN: Asthma in Younger Adults Admission Rates seems to be the same population, same outcome measure; so I'm just trying to figure out how those are harmonized, where the overlap might be. Certainly, it creates a lot of frustration, I find, for providers when they do really well on one measure and not the other one that seem to be the same construct.

MS. BAL: So if you feel that way, we

can definitely reevaluate that when we get to the related and competing. If you want to list additional measures for consideration underneath that, we can definitely do that.

But as I mentioned earlier, we do vote on the measure as-is and then speak about related and competing later. Were there any other questions before we vote?

(No response.)

MS. BAL: All right. So voting is open for overall suitability for endorsement. The options are 1 yes, 2 no. This is your initial recommendation to CSAC and the members.

(Voting.)

MS. BAL: Okay. So for the vote for overall suitability, we have ten yes, 12 no. So we actually are in the grey zone for overall suitability, so when we got out to comment, we write up the draft report, we'll make sure to mention and ask for comments specifically towards this. And since we were in the grey zone for overall suitability, we will revote for this

measure.

So, thank you, everyone. And it is break time. Since we did have that technical difficulty break, perhaps we could do ten minutes instead of five, would that be okay?

CO-CHAIR LANG: Yes.

MS. BAL: All right. Thank you. We'll take a ten minute break. Oh, never mind, we can just go ahead -- yes, we can do 11:00.

(Whereupon, the above-entitled matter went off the record at 10:46 a.m. and resumed at 10:59 a.m.)

CO-CHAIR LANG: We're going to resume. Everyone, please take your seats. There's been a change from the agenda that was distributed in the sequence with which we will address measures in the next segment such that we will address measure 2856 now.

Measure 2856 is entitled Pharmacotherapy Management of COPD Exacerbation. This is an event-based measure that was previously endorsed as number 0549, but the

endorsement was removed in the last pulmonary project in July 2012 and the developer has resubmitted the measure for consideration. The measure -- actually, if the developer -- do we have -- Lindsey, Mary, do you wish to briefly summarize the measure in two to three minutes for us? Thank you.

MS. ROTH: Hello. I'm Lindsey Roth, Senior Healthcare Analyst at NCQA, and this is Mary Barton, Vice President for Performance Measurement. So, just to summarize the measure, this is an episode-based measure that assesses whether patients 40 years or older who appeared in the emergency department or the hospital for a COPD exacerbation are on or were dispensed corticosteroids and bronchodilators.

And two rates are reported. There's the percent that were dispensed a corticosteroid within 14 days of the event and also the percent who were dispensed a bronchodilator within 30 days of the event, and the measure is based on clear guidelines and evidence that administering

corticosteroids and bronchodilators following an exacerbation reduces symptoms and shortens recovery time. And the measure is important to improving quality in this area.

So this is a health plan level measure and it uses administrative and pharmacy data to assess whether medication was dispensed to the patient. And we will be discussing several health plan level measures today that capture medication that was captured through administrative and pharmacy data, which is really the best approach we currently have for measuring medication management at the national health plan level at this point.

But over the past several years, we have had conversations both with the Pharmacy Quality Alliance and with health plan representatives about the potential for patients who may be dispensed medications outside of their pharmacy benefit or at pharmacies such as CVS or Walmart and how this may affect capturing dispensed data that's not paid for through the

patient's health plan. But what we have learned is that health plans are working hard to get this data from pharmacies and even going so far as to provide them incentives in order to get the data and that pharmacies are really willing to feed the data to plans because of their mission for improving quality.

So, because of this improved data exchange, there's been a lot of improvement in health plans' ability to capture prescriptions that were dispensed outside of their health plan. And we also had the opportunity to discuss this measure with the previous steering committee and we did bring the measure back to our expert panel after it lost endorsement to discuss some of the issues that were raised last time.

One of the issues that came up was related to capturing medication samples that are dispensed in the emergency department. And this is a practice that would actually affect all medication related quality measures and there are many NQF endorsed measures that currently do not

capture medication samples. But for this particular health plan level measures, we did survey several ED physicians and they confirmed that samples for the corticosteroids are rarely provided. Samples given in the ED for bronchodilators may be a little bit more likely.

But what we had done is that we specified the bronchodilator rate to account for situations like this; so we actually allow the patient to receive the bronchodilator within 30 days following the event, and this is if they were given a sample or had their prescription filled in the ED that day. The typical day supply is 30 days, so this would allow them to have their prescription refilled in the outpatient setting within 30 days. And so, therefore, the patient would count toward the numerator.

So, we really tried to specify the measure to account for all situations where a patient was already on or received the medication following the event in a timely manner. And,

finally, I just wanted to mention that the measure is a health plan accountability measure that's used in national public reporting programs and that plans have demonstrated around a six percentage point improvement since 2008 on both measured rates and there is further room for improvement, particularly on the corticosteroid rate.

CO-CHAIR LANG: Thank you. And we have -- to discuss this, we have Crystal and Thomas. Take it away.

DR. RILEY: Thank you. I'll start with evidence, I think. So, this measure was based primarily on two clinical practice guidelines that seem to have pretty robust levels of evidence. So the practice guidelines that were utilized encompass well over 600 studies between the two of them, so there seems to be a pretty robust level of evidence that this measure is based on.

I think one of the primary questions that we had when we started reviewing this

measure was why the endorsement was initially pulled and why it's being brought back now? And I think Lindsey just touched on that a little bit, so just a question for the developer; given that there's a note in this guide that the specifications for the measure were largely unchanged from the initial measure that was found under a different NQF number and this current one, was looking at the administration of medication samples the only thing that was reviewed or revised or were there other components that were sort of adjusted that compelled you to bring this measure back under review?

MS. ROTH: Thanks. Thanks for that question. So, the reason the Measure did not pass last time was related to the validity component, which the main issue was discussion around the providing samples of the medications and also I think due to some confusion around how we were defining active medications. So, I think my comments earlier touched more on the ED sample

issue, but those were the two main reasons.

DR. RILEY: Okay. And just looking at the patient population, we're looking at acute inpatient discharge or ED encounter, was there any thought given to patients that might show up at an urgent care facility? Was that lumped into ED or was that looked at separately or just sort of disregarded as an option for patients to receive care?

MS. ROTH: So, we did not look at urgent care. We did separate -- we looked at emergency department and hospital admissions only.

DR. RILEY: Okay.

DR. BARTON: There's still a lot of improvement to happen --

DR. RILEY: Right.

DR. BARTON: -- at those more acute settings, but I agree that it would be interesting in the future, should these top out and everyone's doing great on the ED, then go down a level of acuity and look at what's

happening in urgent care centers.

DR. RILEY: Okay.

CO-CHAIR BRATZLER: So, it's been a while since I've looked at these particular guidelines. I guess what worries me about -- particularly when we get to the gap, but talking about the need for the metric, is that you're looking at corticosteroids prescribed within 14 days, or filled essentially because you're looking at pharmacy data, and bronchodilators within 30 days. So, I would be shocked that many patients with an acute exacerbation of COPD come through an ED and don't get one or both medications.

And so I guess the question is at what point in time, if the patient has a relatively quick and good response to the initial management in the emergency department, is it necessary to continue that? I don't like keeping patients on corticosteroids for any length of time after an acute exacerbation if they don't continue to need them.

So, I guess we'll get more into validity and other things, but how often -- I mean, have we done reviews to look at the cases that are failing the measure to see what's actually happening to them in the emergency department in terms of defining that gap?

MS. ROTH: So, I think what I can say is that we based the measure as close as possible to the guidelines. And so we cited several, the main ones being the GOLD guidelines, that clearly state that the patient should be provided these two medications within a timely manner. And so, I think that's -- we just wanted to align with those as closely as possible. I don't know, Mary, if you want to add anything?

DR. BARTON: I think I would just add that our respiratory advisory panel includes a lot of clinicians who do have direct experience in practice and in health plans and at no time -- and as well as the fact that NCQA uses a feedback loop where we ask people who implement our measures to report to us any difficulty with

reporting the measures or any clinical uncertainty around the specification of the measures, and we have not heard that particular criticism from our user group, from our advisory panels, so while I can't give you a positive saying that I've heard your exact point negated, I've also not heard it brought up.

CO-CHAIR BRATZLER: I guess the other point I wanted to make is you alluded to getting pharmacy data from other sources. So I know what I'm seeing in our practice a lot is patients going and getting their $4 medications, which is a lot less expensive than the co-pays for some of the pharmacy plans that they're dealing with.

So, how much of that data do you actually get? We find it to be very, very common for patients that don't show up in the pharmacy database, but we know the patient's getting the medication because they go pay $4 for it.

DR. BARTON: And that's where the health plan has really better -- more leverage than a physician practice has, because the main

providers of the $4 medicines are huge chains like Walgreens, Walmart, and health plans often work with them to get the data; that's the situation Lindsey spoke of where health plans actually provide incentives to pharmacies to report data to them because they want to have a complete record and the pharmacies are willing to do that. So, I think that's the answer of how we see the health plans performing at the level that we see.

DR. LAMPONE: I had one additional comment on that point. I think when you're dealing with the health plans, one of the challenges they have is being able to identify their membership going to that particular chain. So, if that member does not present their insurance card and they're getting a corticosteroid filled, which even off-plan is very, very cheap, for example, prednisone. Maybe some of the short-acting inhalers are a bit more expensive, but they may be able to get them cheaper at other chains or even the VA, a lot

less expensive.

So if they don't present their health card, the health plan doesn't know that they are there. And, as you look across the states, there's different levels of penetration, if you will, in engaging the national chains at the state level and at the plan level in those states to get that information back. And it does cause some burden of data collection on the plans, so they have to do a lot of work. But you're absolutely right, there is initiatives underway to close that data gap with the plans.

CO-CHAIR LANG: Are there any additional comments by members of the committee or on the phone? Questions for the developers? Okay. We're going to vote on the evidence.

MS. AMIRAULT: So, we're going to go ahead and vote on evidence for measure 2856. Your options are 1 high, 2 moderate, 3 low, and 4 insufficient. Just one second, sorry about that. Okay, polling is open.

(Voting.)

DR. NISHIMI: David, we're still waiting for your vote. So, can everyone in the room please point to Janine again and vote. And David, we're still waiting for your vote.

(Voting.)

DR. NISHIMI: Okay, we got it, David.

MS. AMIRAULT: Okay. So there are nine votes for high, 12 moderate, one low, and zero for insufficient, for evidence for 2856. And based on the percentages, we can move along.

CO-CHAIR LANG: Thank you. Crystal, Tom, performance gap?

DR. RILEY: So, looking at performance gaps, the developer notes that COPD does have a rather high incidence, 15 million adults impacted each year, presumably high cost; and there's also a predicted rise in mortality rate over the next several years, about 30 percent. So there is an opportunity to engage early and sort of head the incidence off.

They've also indicated that there are disparities found across several aspects of the

patient population; looking at gender, racial, ethnic, area, socioeconomic status, looking at payers, and also patients with various comorbidities. So there is a high occurrence of disparities in COPD as well that should be looked at.

DR. LAMPONE: I would agree. I have no other additional comments to that. It's interesting that the data that they did develop -- that they did report out showed that Caucasians had a 45 percent higher mortality rate than other races or ethnicities, although African Americans had a higher ED visit, which brings up the issue of access to care and are there problems in that domain, either to medication or to proper follow-up and adherence?

DR. RILEY: As well as variance in the services that they actually do receive when they do present to the ER or inpatient setting.

CO-CHAIR LANG: Thank you. Is there any other additional discussion questions?

DR. DiGIOVINE: I just had a question

about the data that you show around use of systemic corticosteroids and bronchodilators from HEDIS. Is that data after an exacerbation or is that just data in patients with COPD? And the second is, I'm not sure I'm understanding, are there data that show there's a difference by socioeconomic background for this measure across populations?

MS. ROTH: So, I can speak to your first question. So, the data that we presented from HEDIS is based on patients who had an ED visit or a hospital admission with COPD as the primary diagnosis. So it's based on those who had an exacerbation and were hospitalized or went to the ED for that. And your second question was related to -- could you repeat that again?

DR. DiGIOVINE: Yes. Just the second part is -- so using that data or any other data, I didn't see a gap in performance across socioeconomic groups on this measure. So, you give like 77 percent of HEDIS patients have a corticosteroid prescribed, does that differ based

on race or socioeconomic status?

MS. ROTH: Okay. So for HEDIS, we actually -- we don't collect data on socioeconomic status. We can sort of get at that by looking at the rates for Medicaid plans compared to the commercial or Medicare plans, but -- Mary, do you want to say anything about the disparities?

DR. BARTON: Well, just that the Medicaid plans are somewhat lower on the corticosteroid rate, not necessarily on the bronchodilator rate, but on both, they actually have a wider variability, which means that the distance between the tenth percentile and the 90th percentile plan is larger, which means that there's still a bunch of plans that are performing considerably lower than this median.

DR. DiGIOVINE: Can I -- I just wanted to follow up on that. Because this comes up, I think, for a number of measures and so I'm trying to understand how we assess what is normal variation across a population and what is

disparate care.

So, we wouldn't expect that every plan is going to have 70 percent compliance if that's the mean, so what is variation that somehow statistically we decide is more than what we're going to see because we're humans and there's going to be a normal distribution?

DR. BARTON: I think when there is a distribution around, say, the commercial rate for bronchodilators that shows that there's a 16 point difference between the tenth percentile and the 90th percentile, then you have a sense of how closely grouped around their median performance, which is, I think, 75 or 81. Right? So you have roughly eight points on either side of 81 that includes everybody from the tenth to the 90th percentile. So that's a huge majority of plans.

And for Medicaid plans, in contrast, around a rate of 66 for corticosteroids, they've got a 20 -- is that 20? 30 percent difference between the tenth and 90th. So you have a wider distribution, which means you have more people

further down, and I would suggest that a rate of 40 percent on getting appropriate medications to someone with a COPD exacerbation should not be thought of as a random variation, we're all human, but as a quality problem.

DR. DORMAN: Point of clarification, since it's the 30 percent and it's on either side of the mean or the median, then it doesn't go down to 40 percent. It's only 15 below 75, so the range would be 60 to 90 for the 30 percent.

CO-CHAIR LANG: Are there additional questions for the developers or additional comments, members of the committee? All right.

MS. AMIRAULT: Okay. So we can go ahead and vote for performance gap for measure 2856. And the options are 1 high, 2 moderate, 3 low, and 4 insufficient. Again, for performance gap for 2856.

(Voting.)

MS. BAL: Could everybody in the room vote one more time? No, never mind, we're good. Thank you.

MS. AMIRAULT: Okay. So, 13 for high, seven moderate, two low, and zero insufficient. And based on the percentages, we can move along.

CO-CHAIR LANG: Thank you.

Reliability?

DR. LAMPONE: A couple of things that came to light when looking at reliability; during this reliability testing and performing the beta-binomial measure scores, it looks as though that when they had updated data submission, the scores -- there was some discorrelation between the scores. And if you look at the bronchodilator indicator for 2015 in the commercial population -- and also the systemic corticosteroid correlations -- those went down in both the commercial populations.

So, I was wondering again whether that was a data issue? The correlations for Medicaid and Medicare were pretty much in line and I wonder if there's any -- had you had any discussion around that indicator and why the difference in results in the updated data set?

DR. BARTON: Thank you. The beta-binomial looks at data that's reported from plans to try and distinguish to what degree the differences between plans are legitimate reflections of quality problems. And I will say that the beta-binomial works the best when there's a wide spread of responses or of indications.

So when we have the relatively narrow band between the tenth and the 90th percentile, and so we have everybody performing relatively tight band around -- whatever it is, 77 or 80, then the beta-binomial doesn't have very much difference to base its -- to power the mathematical calculation. So, we often see when there's a relatively well grouped rate that the beta-binomial is not as high as when there's a more varied distribution. That's just kind of a mathematical consequence of using the beta-binomial.

DR. LAMPONE: The other scores seem to be in line and I have no further comments on the

reliability test.

CO-CHAIR LANG: Okay. Thank you.

DR. O'BRIEN: Just a question regarding the measure specification, how it's handled if a patient is on -- either has these medications at home and is increasing the dosage or has them on-hand and then just resumes taking them?

MS. ROTH: So, this is related to the measure counting patients who have an active prescription for the medication and so what a health plan does when they're calculating this measure is they identify the person in the denominator and then they look back to see if the patient had an active prescription for the bronchodilator, the date it was dispensed to them, and then the day supply.

And so, as I mentioned before, the typical day supply is around 30 days. So, it's not like the health plan is having to go all the way back a year ago. They're really looking a short period of time back to see, were they dispensed the medication on this date? And so,

if the day supply is 30 days and the patient was hospitalized 15 days after that, then they would be in the numerator because they had a prescription filled for that medication already and were already theoretically on -- taking the medication.

DR. O'BRIEN: So, just to clarify, the 30 days is the cut point? If they filled that prescription 30 days prior to that, they would be considered then a patient that had a miss?

MS. ROTH: Yes. So if it's --

DR. O'BRIEN: Okay.

MS. ROTH: -- 30 days or more, then yes, they would not be in the numerator.

DR. DiGIOVINE: I just want to clarify a point as well. If the data is being used to compare health plans, so am I right in assuming that -- the majority of patients that are going to be in there are patients with commercial insurance where the reliability is the least well supported?

DR. BARTON: The HEDIS looks at

Medicare plans, commercial plans, and Medicaid plans. And it's true that the population is higher for this measure in the Medicare plans, but the commercial plans that report have met a threshold of number of patients who meet the criteria.

So, there are no plans that are reporting on very small numbers, but it's true that the average number of people in the measure is going to be higher in the Medicare plan than in the commercial plans.

CO-CHAIR LANG: Are there any additional comments, questions for the developers?

MS. AMIRAULT: You can go ahead and vote on reliability for measure 2856. And the options are 1, high, 2, moderate, 3, low, and 4, insufficient. Again, for reliability for measure 2856.

MS. BAL: Ella and David, if you're still on the call, please vote. Operator, could you also check David's connection, he said he was

having difficulty. Oh, never mind. Never mind.

MS. AMIRAULT: So we have five for high, 16 moderate, one low, and zero insufficient. And we can move forward based on the percentages.

CO-CHAIR LANG: Thank you. Validity?

DR. LAMPONE: Okay. We'll move on to validity testing. They underwent construct validity testing and that was new since the prior submission. They also included in validity testing sensitivity and specificity, which was also new. We did have a question on the data elements for the validity testing using the sensitivity and specificity. It appears as though you only used four plans worth of data, is that right? And in the total, you had much higher participation in plans. When you look under construct validity, you quote 241 commercial plans, 357 Medicare, and 157 Medicaid, was that just an added sample just to test to see the specificity and sensitivity?

MS. ROTH: So, there are two different

ways that we had tested the measure. There was the original field test data that we had conducted in 2004 and there's also more up-to-date HEDIS data that we had provided from, I believe, year 2014. And so, that's why you're seeing the differences in the numbers there.

DR. LAMPONE: Okay. And --

MS. ROTH: Right. And we, I just wanted to clarify one thing is, we actually did not conduct the sensitivity/specificity analysis that was on Page 7 of the worksheet. It was based on our field test data from 2004. But that was something that NQF had conducted --

DR. BARTON: Right. We took the field test data, as the PA indicates, and NQF Staff converted that to those values.

DR. LAMPONE: Okay. Now for me, the sensitivity was a little bit low, but the specificity being higher I think drove the 88 or 85 percent positive predictive value, so that made me feel a little bit better. And then, on the Pearson correlation coefficients, they were a

wide range. Can you comment on that? Is that, the wide range, is that based on the different percentile performances? So, from the tenth to the 90th? That's on the bottom of Page 7.

MS. ROTH: So for the construct validity testing, the good thing is that we did look at the correlation between the two rates within the measure and those were highly correlated, which is --

DR. LAMPONE: Okay.

MS. ROTH: -- what we would expect. We also did conduct correlations between this measure and another related, similar measure for looking at spirometry testing following a COPD diagnosis, so it's not -- the connection between how those two measures are related is a little bit less than the correlation between the two rates within the measure, but it is something that we wanted to look at and provide a data for, so the correlations for that are not quite as strong.

DR. LAMPONE: Okay.

DR. DiGIOVINE: Sorry about that. I'm really struck by your field tests and I'm just trying to make sure I understand it right, what I'm understanding you to say, when you looked at all of these cases that about 50 percent of the time, your administrative data is saying the patient has a COPD exacerbation, there is no evidence in the medical record that that actually occurred. Is that correct?

DR. BARTON: So you're talking about the data about confirming exacerbations or data about filling medicines?

DR. DiGIOVINE: Yes, the first one is, I'll get to the medicine second, but the --

DR. BARTON: Okay.

DR. DiGIOVINE: -- first table that you have is the COPD exacerbation medical record validation by plan and product line. It's not numbered, it's the first one under the results of critical data element validity. It's on the bottom of Page 38, start of Page 39.

DR. BARTON: Okay.

DR. DiGIOVINE: At least in our final version, which may not be the one you have.

DR. BARTON: Yes. Well, so, I guess the generic point that I want to make about the 2004 field test that used a -- where we asked a small number of health plans to use all the data that they had at their disposal. So, administrative data, pharmacy data, and pulling medical records. It does not surprise me that there are events that happen that patients have a claim for that don't get reflected in their medical record.

It could be that the health plan couldn't find the relevant provider who was taking care of that patient for that problem. Maybe they found, you know, a primary care chart and the patient had a pulmonologist, they didn't find the pulmonologist chart. I mean, I think that the completion of the data we can be sure about is much higher for the claims and the pharmacy claims. If you get in an emergency department, that hospital will bill your insurer

100 percent of the time.

And so, the data for looking at which data can we trust, there are places where the medical record either does not reflect something that we could say with a pretty high degree of confidence probably happened and then conversely, there are things that show up in the medical record, because I'm presaging the medicine points, there will be doctors who will write, patient got X and Y meds, but the patient never filled them.

And so, it appears in the medical record that they have been prescribed a medicine, but the effective result is apparent from the pharmacy claims that they never picked it up. So, there is -- we would not have presented the sensitivity and specificity data in the way that it was presented to the steering committee because we think that there, again, looking at that 2004 data, there are enough questions in my mind about the completion of the medical record data and the accuracy of the medical record data.

MS. ROTH: I would also add quickly that I think we did note in here as well that one of the limitations of this testing that we had done was that none of the four plans had access to any medical records on the hospital side. And so, that's another reason why the COPD diagnosis might be missing in the medical record in this data.

DR. DiGIOVINE: So, just to -- is it -- I mean, I think it's possible that somebody could go to an emergency room, you could have record of that, but on a medical record review, you decide it's not a COPD exacerbation, it's a pneumonia, it's heart failure, and it somehow got coded as a COPD exacerbation. Do you have any sense of to what extent that was the issue?

DR. LAMPONE: How about your specificity data? Would that address that? You have --

DR. BARTON: I'm sorry --- so, you have to have COPD to be in the denominator of the measure, so I --

DR. DiGIOVINE: I understand that, but you can go to an emergency room for something other than a COPD exacerbation.

DR. BARTON: Of course. And so, you're saying that the percent of time that the ED is wrong in how they assign the diagnosis, it's not actually a COPD exacerbation, and I don't have data on that.

DR. DiGIOVINE: Okay. And then, so in terms of the second part, again, your data is saying that somewhere between 30 percent of the patients that you're saying do not get corticosteroid or a bronchodilator, in fact in the medical record there's evidence that they were prescribed those medications? Does that strike you as problematic in terms of the validity of this measure?

DR. BARTON: Not when so many plans have been able to show a rate of 80 percent compliance. So, it does not -- if the improvement went from 60 to 80 percent overnight even, you would say, okay, well, they figured out

how to capture the medications that are actually being given. So, if the providers somehow have a way that's not captured of giving medications, health plans have figured out how to capture that. So, I can't really speak to what was going on in 2004 at the time of the field test, I was not part of the organization at that time, and there's actually nobody who's still part of the organization I think who did this field test, so there's probably some unanswered questions that we all have.

DR. DiGIOVINE: Okay.

CO-CHAIR LANG: Thank you. In the interest of time, we're going to move on to vote on validity.

MS. AMIRAULT: So, we'll be voting on validity for measure 2856. And the options are 1, high, 2, moderate, 3, low, and 4, insufficient. Again, for validity for 2856. So for results, one high, 13 moderate, eight low, and zero insufficient. And we can move along.

CO-CHAIR LANG: We're running a few

minutes behind, so if you could comment briefly on feasibility, Crystal, Tom?

DR. LAMPONE: Okay, sure. Feasibility, really this is data obtained from the medical records. I think some of the threats to that is the ED access to medical records. And then also, where you have some impact on the data where you have patients, especially in the Medicare population, who may be on chronic corticosteroids or frequent corticosteroid dosing and certainly inhalers that may impact the feasibility in the sense that it may be overstating that they're already on a medication and then they're compliant and moving on that way. But, I think overall, those are the only concerns to feasibility and those getting prescriptions filled at other places.

CO-CHAIR LANG: Is there anyone else who has any comments, questions for the developers? If not, we will proceed to vote on feasibility.

MS. AMIRAULT: So, we'll be voting on

feasibility for 2856. And the options are 1, high, 2, moderate, 3, low, and 4, insufficient. Again, feasibility for 2856. So the results, two high, 17 moderate, three low, and zero insufficient. And we can move along.

CO-CHAIR LANG: Crystal, comments on usability and use? Tom?

DR. RILEY: Sure. I'll try and make it brief.

CO-CHAIR LANG: Thank you.

DR. RILEY: So, from a usability perspective, it looks as though this measure is already used pretty widely. It's been incorporated in public reporting for NCQA State of the Health Care Annual Report and Quality Compass, used in the Medicare Advantage Star Rating Program, et cetera, et cetera. So from a usability perspective, we didn't identify any issues.

DR. LAMPONE: I agree.

CO-CHAIR LANG: Other comments, questions for the developers? Thank you. We'll

proceed to the vote.

MS. AMIRAULT: For usability and use for 2856, the options are 1, high, 2, moderate, 3, low, and 4, insufficient. Again, usability and use for 2856. So 16 high, six moderate, zero low, zero insufficient and based on the percentages, we can move along. Okay. So, now the overall suitability for endorsement, 1 is yes, and 0 no, for measure 2856. Okay. Seventeen yes and five no.

MS. BAL: Could everyone turn off their mics? Thank you. So, with that, this measure is recommended for endorsement and we can move on to the next measure.

CO-CHAIR LANG: Lindsey, Mary, thank you. So, now we will proceed to address two measures. The measure steward for each of these, they're similar measures, the measure steward is University Hospitals Cleveland Medical Center. And the first is 2794: Rate of Emergency Department Visit Use for Children Managed for Identifiable Asthma, a PQMP measure. The second

is Appropriateness of Emergency Department Visits for Children and Adolescents with Identifiable Asthma, a PQMP measure. We will address these individually, however. And, Larry, would you like to summarize the first measure, taking perhaps two to three minutes?

DR. KLEINMAN: It would be my pleasure.

CO-CHAIR LANG: This is Larry Kleinman, who is with us today.

DR. KLEINMAN: Good morning. I'm Larry Kleinman. I am the measure developer, which really was CAPQuaM, the Collaboration for Advancing Pediatric Quality Measures. What I'd like to do is very briefly give a history of the development of the measure and then say a couple of things about the measure, and partly in response to the call that took place. I'll try to get this all within two to three minutes.

First of all, CAPQuaM is one of seven Centers of Excellence that emerged from a peer review competition held by AHRQ and CMS to create the Pediatric Quality Measures Program, something

that was defined and created legislatively as a part of the CHIPRA, the Children's Health Insurance Program Reauthorization Act of 2009. Our tasks were to improve and strengthen children's healthcare quality measures, expand on existing pediatric quality measures, and advance their development in order to increase the portfolio of quality measures available to public and private purchasers of children's healthcare services, providers, and consumers. And that is directly from the legislative act.

For this work, we were guided by the Institute of Medicine definition of quality, which is the degree to which health services for individuals and populations increased the likelihood of desired health outcomes and are consistent with current professional knowledge. And as a part of that, I'd like to point out that this suggests that quality may be a continuum and not only a dichotomy of good and bad. And the various components of quality according to the IOM are timely, safe, effective, efficient,

patient-centered, and equitable care.

We had a very -- our process was peer reviewed as a part of the competition for the grant. It was a highly engaged and transparent process and the various definitions and parameters, I know there were questions of how we got age groupings and things, came from a national expert panel. Our partners in this work included the Academy of Pediatrics, the American Academy of Family Physicians, and NCQA was a partner, New York State Medicaid was a partner, there were a large national partnership. I'm happy to give more of that if there's questions.

The panel used a RAND style method. And the panel not only created the parameters, but once we created the measure, they, without objection, endorsed the measure. A couple of things that relate to the measure itself is the rate of emergency department visits for children with identifiable asthma, defined as visits per 100 child years. We discussed -- so, I can stop there and we'll take it wherever you'd like it to

be.

CO-CHAIR LANG: Okay. Thank you, Larry. The discussants for this measure are Don and Kim. Please begin with a discussion of the evidence.

DR. ELLIOTT: I'm happy to. Don, I think, is on the phone, so he can add to anything that I miss. But the measure number was already talked about, 2794, and the measure title is Rate of Emergency Department Visit Use for Children Managed for Identifiable Asthma. It is a PQMP measure. The description of the measure estimates the emergency department visits for children ages two to 21 who are being managed for identifiable asthma. And it is a little bit different measure method than I'm used to seeing, which is the rate of visits per 100 child years.

The level of analysis was done at the health plan, the integrated delivery system population community and it can be broken down or stratified by county or city population and other various stratifications. The evidence for the

measure indicates that it's an accessible, high quality, primary care reduces the amount of ED visits or the frequency of ED visits related to asthma. And a systematic review of the body of evidence was not required because this is an outcomes measure. However, the developer did provide evidence that based on the NHLBI Guidelines from 2006, there is evidence to support the measure, as well as they did a pretty comprehensive literature review, anywhere from 2004 through 2006, and they reviewed 4,747 abstracts. So, the literature review was pretty comprehensive. Don, do you have anything to add?

CO-CHAIR LANG: Don unfortunately is not on the phone right now. Basically, the premise here is that ED visits can be avoided if asthma management is improved strictly in the primary care setting? Is that correct?

DR. KLEINMAN: Yes.

CO-CHAIR LANG: And there is a difference from an epidemiologic standpoint in terms of how you've set this up, if you could

please elaborate briefly on that for members of the committee?

DR. KLEINMAN: Sure. We were looking to enhance an existing measure that was a part of the CHIPRA Core Set. That measure looked at the percentage of children with an asthma ED visit, but you could qualify for that measure simply by having an asthma ED visit and never being known. We felt a more accurate measure would be represented by requiring identification of the child as asthma to proceed the ED visit in order to make it something that someone might be held accountable for.

In order to do this and appropriately weight the influence of any individual ED visit, we recognized it was more accurate to do so by having an actual rate and using a person timed denominator. So, if the child became identifiable according to the definition six months into the year, what happened in the first six months would not contribute to the numerator, only what happened in months seven through 12.

And so, we felt this was a more valid, more useful representation, and that because it was an epidemiological definition, it actually was pretty readily understood with just this basic explanation.

CO-CHAIR LANG: Thank you. Are there additional questions for the developer or comments from members of the committee?

DR. DiGIOVINE: Sorry, just a quick question. Are we measuring physicians in this or health plans or both?

DR. KLEINMAN: Not measuring physicians, not measuring the hospitals that they're going to, measuring health plans and up.

CO-CHAIR LANG: We'll proceed with the vote on the evidence.

MS. AMIRAULT: Okay. So voting on evidence, health outcome, 1 is yes and 2 is no for measure 2794.

DR. NISHIMI: Can everyone point again? We're only looking for 21.

MS. AMIRAULT: Okay. So 21 yes, zero

no, and we can move along.

CO-CHAIR LANG: Kim, performance gap?

DR. ELLIOTT: Yes. I think that they did a really good job, the measure developer did a pretty good job on performance gaps. It does clearly indicate that there are some disparities, particularly in two different age populations, the very young and the adolescent. It also showed a pretty good stratification in those age groups. I think some of the other things that they pointed out is that there are some disparities between rural and urban, so clearly there are some opportunities for improvement. And Hispanic and black, also those populations had some disparities. So, from an opportunity to really make a difference by measuring the outcomes, I think the developer did a good job identifying that there was opportunity for improvement.

CO-CHAIR LANG: Dale?

CO-CHAIR BRATZLER: So, I completely agree, there's opportunity for improvement, but

the question is, is that a true gap? I mean, so I'm not a pediatrician, so I don't know whether we expect asthma admission rates be different by different age groups, I don't know. So that may be one of the reasons that there's a quote, disparity by age group, and then we have all those issues around sociodemographics that clearly impact rural versus urban, black versus white. I mean, so definitely disparities, but is that a gap in the actual care or does that reflect the system that prevents supportive care? And then, were there differences by age group?

DR. KLEINMAN: In general, the disease is not considered to be biologically different by age group, but clearly there are utilization patterns that differ. I think regarding some of the other demographics and I'd done this in a number of areas, but I brought the data simply for black versus not black because I think that the black was the highest utilization rate so it allows it. If you look at black versus --

MS. AMIRAULT: Sorry. Larry, you're mic

needs to be on.

DR. KLEINMAN: Okay. It was lit there but not here. Sorry. So, what I'm saying is that there are not fundamental biological issues that are recognized by age, but there definitely are utilization issues. So, age is important probably to consider. In terms of black and white, I ran the data after the phone call by plan and race to look at the difference. And if you look at it, between black versus not black, there's a mean difference of about ten units, ten visits per 100 child years.

But there are plans for which that gap is 2, 2.2, 2.4, 2.8 on the low end and others at 15, 18, 15.6 at the high end, suggesting that this is not determined solely by demographics, but that really plans seem to matter. And these differences were significant and meaningful and the distribution, it was pretty dramatic to see it on the little graph.

CO-CHAIR LANG: May I ask you a question, follow-up to that right away? You

said, not solely determined by demographics, your data implies that it's not determined solely by race? I mean --

DR. KLEINMAN: I'm sorry, yes.

CO-CHAIR LANG: -- in order to say that, you would need to adjust for other -- were other sociodemographics factors included in your analysis?

DR. KLEINMAN: I'm sorry, before I had the mic on, I had mentioned that I did it for others such as rural versus urban. I didn't bring those data with me. The differences were not as dramatic and, therefore, the data were not as dramatic, but the pattern was the same.

CO-CHAIR LANG: Let me rephrase the question. The other demographic factors that would be important in your analysis in addition to race would be, for instance, zip code of residence, income, educational attainment of parents, et cetera, which also as we all know can influence outcomes.

DR. KLEINMAN: What I would say is the

NHLBI Guideline says that this is attainable in all children. And if we're talking about at the plan level, as opposed to at the practice level, which is where this is, plans contract to manage with the populations that they have. So, we did look -- I did not do this same breakdown between the phone call and today, but I do have data looking at poverty at county of residence, I don't have individual level. These were all Medicaid patients, by the way. So, this was within already a difficult to manage population, but we see significant broad variation. And we also see across plans that there's a large range of performance rates and it's not solely dependent upon whether, you know, they're black or poor, it's determined by other things.

CO-CHAIR LANG: So then, again, to follow up, your data show that there's a difference by race, not -- so, you broadened it to demographics, which is a much broader category, as you know, that includes other factors. So, I'm just trying to be precise and

understand what you're saying.

DR. KLEINMAN: Race and urbanicity is what I have examined. And I would say that -- well, I think that the data suggest that populations that are historically thought of as hard to manage, some plans do much better than others. And the ideal rate for this, I should say, because this also comes to the rural/urban consideration, is not zero. You would expect that there are some children who have had terrible deteriorations quickly in the past, they go to the ER rather than going to the doctor.

There are exacerbations, there are children who are well managed who have really bad asthma or who break through, who are exposed to something or the season changes or the weather changes. So this is not -- this is why it is really granted as a continuum. And if these data are interpreted with a good deal of thought, the higher numbers in urbanicity relate probably both to management and availability of emergency rooms in contrast to at least the rural areas. So

there is an access component that probably plays into this.

CO-CHAIR LANG: You just -- so we're going to be voting on performance gap momentarily, but again, some of these issues are crossing over more into validity as opposed to performance gap. So we need to vote on performance gap now. Thank you.

MS. AMIRAULT: Okay. So voting for performance gap for measure 2794, 1, high, 2, moderate, 3, low, and 4, insufficient. So for results, six high, 14 moderate, one low, and zero insufficient. And based on the percentages, we can move along.

CO-CHAIR LANG: Kim?

DR. ELLIOTT: Okay. Let me move into the reliability. And there was some concern from Don regarding the reliability that he sent me and that was basically that it really is hard to attribute the ED visits to really being an outcomes measure for what they're trying to represent it as. And that was really the only

comment that he wanted me to specifically make. But, as far as the information that the developer provided, the results are not risk adjusted, so it is a pretty straightforward measures.

One of the concerns that I had was related to the pharmacy data and possibly some of the pharmacy data not being available to really determine the outcomes for this particular measure and that the developer indicated that the pharmacy data really wasn't critical, yet asthma is clearly a pharmacy driven measure, even coming out of an ED. So, that was a little bit of something that I think the developer really needs to address with us as well.

DR. KLEINMAN: Okay. Thank you. One thing I want to say is in pediatrics, this is a pretty well accepted outcome measure for performance and it's used in various not consistently specified ways. But in terms of the pharmacy data, I would say that data is not fundamentally critical to the measure. And the reason I say that, and I regret I do not have the

data because the person who did the analysis at New York State is gone and we've not been able to find the analysis, but we found that use of the pharmacy data qualified a few more children into the measure, not a large percentage, something on the order of one to two percent more children were qualified as having identifiable asthma, but did not have impact on the rate.

The pharmacy data was included because we were looking to try to be as inclusive as possible and that was the judgment of the expert panel who felt that the pharmacy data ought to count as part of this. But I would say it is not fundamental to the measure, I think at the margins, it makes it a better measure.

CO-CHAIR LANG: Is there additional discussion? Questions?

DR. ELLIOTT: The only other thing I would point out is that the developer didn't do any additional testing for validity, it relied on previous evidence.

CO-CHAIR LANG: Larry?

DR. KLEINMAN: I believe, and we may not have put this in the application because we're still new at this to some extent, but I believe that the expert panel endorsement of the measure is considered to speak to its validity as well.

CO-CHAIR LANG: We're on reliability.

DR. KLEINMAN: Yes.

CO-CHAIR LANG: Okay. So, we're going to vote on reliability.

MS. AMIRAULT: So, for reliability for measure 2794, the options are 1, high, 2, moderate, 3, low, and 4, insufficient. Again, reliability for 2794.

DR. NISHIMI: And we need everyone to point their clicker again.

MS. AMIRAULT: Okay. For results, two high, 17 moderate, two low, and zero insufficient. And we can move along.

CO-CHAIR LANG: Validity, Kim?

DR. ELLIOTT: Validity specifications, the developer relied a lot on literature to

support its conclusion on the validity and they did use administrative data as a primary source. It is at the data element level. The developer did cite face validity, but did not specifically assess face validity at the computed measure score level, and that is something that NQF really does look for. They do have some threats to the validity as well that they talked about.

They did exclude COPD as a diagnosis, cystic fibrosis and emphysema, which is appropriate. And the exclusions are clinical and represent construct validity, rather than statistical validity and considerations. And they did not do any risk adjustment, so disparities are -- they really focus on the disparities rather than risk adjusting the population. And the specifications that they put into the measure did stratify by age group and race/ethnicity, which also seemed appropriate. And they didn't note that there is some variability when they did these stratifications by rural versus urban and at the county level.

Since they used administrative data for their research and their results, there really wasn't an issue with the completeness of the data set. And I believe that's all the comments I had.

DR. GLOMB: My comment, I guess, is this the right time to talk about validity? The comment was that there was not a difference, someone actually might have already said it, that there wasn't really a difference in asthma across the age spectra. And, that really kind of goes against one of the inherent built-ins in pediatric pulmonology. Fernando Martinez, back in the 1990s, published that there really are three different types of wheezers as children move from infancy through at least their first six plus years of life.

And my concern with using emergency department visit as a measure of whether or not the patient is overall being well controlled does in fact exclude viral seasons, et cetera, that are going to adversely affect particularly the young wheezer, the early transient wheezer who

won't turn out to have asthma, the non-atopic wheezer who is going to respond to every viral illness with wheezing. And I guess the expectation is, well, during the viral season, they should all be on ICS, but could you address that please?

DR. KLEINMAN: So, first of all, one of the things we did is that the specifications start from age two up, so the very early wheezers we're taking out of it, the tracheo-laryngomalacia, who are stridorous wheezing, whatever the heck it sounds like to people. So part of it is the definition of the population. The other thing is this notion of identifiable asthma puts a hurdle of prior diagnosis so that if they haven't been considered to be an asthmatic or haven't been managed for asthma within the previous, it could be anywhere from one year to two years, because other way we'd do the assessment on a month basis, I'm sorry, from 12 to 23 months, then they wouldn't qualify. We tried to create enough of a hurdle to avoid at

least a good chunk of those that you're talking about. I share your concern otherwise.

DR. GLOMB: And I -- you know, not to disparage the ER docs, but a lot of times these running diagnoses of asthma come out of their ER visits rather than their routine PCP or other specialist visits. So, if we're relying on a previous ER visit where they were wheezing and the diagnosis of asthma was made to establish the underlying diagnosis of asthma, I --

DR. KLEINMAN: It would be at least two events, not just -- this has to be the third event if they're under six. Over six, it could be the second event.

DR. GLOMB: Thank you for clarifying.

DR. KLEINMAN: And so that's -- but I share your concern, I'm sure there's some noise from that, but I think we've mediated it and moderated that pretty well through the specifications.

CO-CHAIR LANG: Yes. I'm wondering whether this would be the area to discuss the

risk adjustment strategy here or the lack of risk adjustment.

DR. KLEINMAN: Our feeling was that you've got multiple players in this system and we have specifications for -- we ask for it to be stratified by age and race, we have specifications for rurality, we have specifications related to poverty of the home county. And we make all of them available. We didn't require them, because we were trying to balance between burden and precision.

But our -- the way we articulate this is that the accountability entity, the people who are asking for the measure to be used, who after all are theoretically the payers or the regulators, ought to be the ones who make the decisions how many slices into the stratification there is. This will be a more precise measure, assuming that there's adequate sample size, the more it is stratified.

And we tried to provide the tools to do that, but we don't want to exclude plans that

are a little bit smaller or accountability entities from saying, we're contracting with you to manage a population and then of course that ought to be figured into the cost and the way they pay. So, to us this is something that isn't for the measure developer to decide, but is actually a part of the relationship between the accountability entity and the accountable entity.

DR. DiGIOVINE: I'd just like to kind of follow up on that. So, is there a sense from the developer, from you, whether the differences -- to what extent the differences we see across race or urbanity are differences to the inherent quality of the primary care and how much of them are due to environmental factors and social factors that are not under the control of the health plan?

DR. KLEINMAN: What I can tell you, I can speculate to that. I think it's a mix of both. I will tell you, when I did work in Massachusetts Medicaid, I saw, you know, everybody was well above the national average, so

clearly there was something about the way they managed and I don't think the environment is good throughout Massachusetts.

I will also tell you that in this work, if I go back to the analysis I referred to earlier of just looking at black versus non-black, because I thought that was the most dramatic in terms of its potential impact, we see actually nine-fold differences, from 18 to 2 absolute difference in rate between -- no, from one plan to the next, one plan had a gap of 2 between blacks and whites or blacks and all others, and another had a difference of 18. And then if you look at the second one, to avoid the most dramatic, it's 2.2 and 15.6, still a seven-fold. And then, 2.4 and 15, still a six-fold difference. So, I don't think that's all environmental or all demographic, I think that reflects that there are real differences in management happening.

DR. LAMPONE: Across the plans, was the demographic distribution, race in this case,

about equal or did you have some plans that had a higher concentration of African Americans that they may have seen people, because they had a, you know, bigger cohort to follow up on and a more consistent application of their plan policies?

DR. KLEINMAN: I removed from this analysis all low numbers. I don't know exactly the answer. So this is not spurious because of a low number, whether -- so, I don't know. I have the data, I could probably answer that during the break if we needed to, but what I would say is that it's clear that there are skills and ways to manage the black population to make it closer to the non-black population, and in fact close to the non-black population, probably not equal, at least that I can demonstrate in the data. But, again, we provide for stratification because we think it's important that plans know. We also think that the gap itself is an intrinsic measure, the IOM asks equity as a part of it, if you risk adjust, it's much harder to see how

equity comes into play, if you stratify, it's in your face, frankly.

CO-CHAIR LANG: One final question and then we'll move to the vote.

CO-CHAIR BRATZLER: Okay. It's not really a question, it's just a point that David made earlier, that you've clearly shown stratification by race, but I can tell you that race varies dramatically, the impact of race, by location, site, all those other demographic factors that are very important that aren't addressed in this measure or the next one.

DR. KLEINMAN: If I could, we have specifications and we were advised during the process not to put them all in. For percentage of poverty in the county and the rurality, we don't have always easily available patient level data of this. And this was all within Medicaid that we did that, but we also looked at different categories within Medicaid and found actually surprising little differences there.

CO-CHAIR LANG: We will move to a vote

on validity and I want you to note that according to the NQF algorithm for validity, the highest eligible rating for this measure for validity is moderate. So, we're voting 2, 3, or 4, for moderate, low, or insufficient.

DR. NISHIMI: Just to make the reason clear is because to be eligible for the high, you have to provide empirical testing at the score level. And, obviously, moderate is a passing grade. So 2, 3, and 4.

DR. KLEINMAN: And just -- this is where I'll make that point I made earlier that I don't know that it was articulated, but this was approved without dissent by the expert panel as a measure, not only the components.

MS. AMIRAULT: Okay. So, if you haven't voted for validity for 2794, the options are 2, moderate, 3, low, and 4, insufficient. So, zero for high, ten for moderate, 11 low, and zero insufficient. So this will be in the grey zone.

CO-CHAIR LANG: Feasibility, Kim?

DR. ELLIOTT: Yes. Feasibility, it is all administrative data. The fields that they would collect the data from are in the electronic as well as paper based. So, I didn't see any issues with feasibility for this particular measure.

CO-CHAIR LANG: So there are no concerns identified. Anyone else have any comments, questions? We will proceed then to vote on feasibility.

MS. AMIRAULT: Okay. So, feasibility for 2794, 1, high, 2, moderate, 3, low, and 4, insufficient, for feasibility for 2794. Okay. So there are 15 high, six moderate, zero low, and zero insufficient. And we can go ahead and move on.

CO-CHAIR LANG: Excellent. Usability and use, this is a new measure not currently in use. Kim?

DR. ELLIOTT: That's correct. And currently it's not publically reported, it's not used as an accountability for any particular

program, as least as referenced by the developer. It is planned to be used, but there wasn't a lot of information as to who or what would use the measure. They did not indicate any potential harms and there was no additional feedback given on this particular measure.

CO-CHAIR LANG: So, ED visits are already commonly measured by a number of plans around dashboards, what does this add?

DR. KLEINMAN: This adds a number of things. First of all, it's the concept of identifiable asthma is a new construct that, when we validated it against things like prevalence, it was about half the prevalence, we validated it against something like persistent asthma, which was a much tighter diagnosis, the NCQA diagnosis was about two and a half times as many it identified, so it achieved that. We also recognize, we looked at New York State Medicaid data and also national data and found that when you look at emergency department visits through billing data, if the child is hospitalized, most

of the time those data are not actually forwarded.

So when we defined -- and we found that somewhere between 70 and 85 percent, in New York State it was 83 percent, of admissions for children with asthma come through the ED. So, if you actually want to accurately estimate the number of emergency department visits, the best way to do that is not simply with ED visits, if you're using this data, but with ED visits and hospitalizations, and then deduplicate. So, that's something else that we add. It's a more accurate assessment than other sorts of data. It's also specified for children and the rate, the fact that it uses a person timed denominator actually introduces a new construct into accountability measurement that I think will be valuable for other things as well.

DR. GLOMB: Just before we leave this, and under unintended consequences and I don't want to belabor the point any more than I have to, but I still am fairly concerned about lack of

stratification by risk here. I understand what you're saying, I understand what your panel felt, I'm going to take our real world experience across Texas, not doing some sort of risk adjustment on the patient base may give some very, very skewed results.

Health plans that are predominately in the urban areas, and I'll take Houston as an example, where their Medicaid plan, where their population might be 60 percent black in an urban area, which rates as the highest asthma risk category, versus a company that is in the predominately the rural areas of this state, where the black population may be only 15 to 20 percent in Medicaid there and where, interestingly enough, the risk to the black patient in the rural area drops almost to the level of the Caucasian patient in those same areas. I think that then doing a cross comparison across the state without taking risk into account would adversely affect that predominately urban plan versus a plan with a

broader base of numbers.

DR. KLEINMAN: Well, we --

DR. NISHIMI: It's not at a point now where we need to discuss back and forth, we can just -- the committee has heard your view, the committee has heard the other views of the other committee members, we just need to vote.

DR. DiGIOVINE: Just as a quick -- this is not a new measure, so I'm just curious as to, have you approached groups to --

DR. NISHIMI: This is a new measure.

DR. DiGIOVINE: Oh, I'm sorry.

DR. NISHIMI: This is a new measure to NQF.

CO-CHAIR LANG: We will proceed to the vote then.

MS. AMIRAULT: Okay. So --

CO-CHAIR LANG: This is -- please go ahead.

MS. AMIRAULT: Oh, no, go ahead.

CO-CHAIR LANG: No, go ahead. Go ahead.

MS. AMIRAULT: Usability and use for measure 2794, the options are 1, high, 2, moderate, 3, low, and 4, insufficient. Again, usability and use for 2794.

MS. BAL: Ella, please vote, we're still waiting for yours.

MS. AMIRAULT: Okay. So, four high, 11 moderate, five low, and one insufficient. And we can move along. So, now we'll look at overall suitability for endorsement, 1 for yes and 2 for no, for measure 2794.

MS. BAL: Ella, we're still waiting for your response to overall suitability, please.

MS. AMIRAULT: Okay. Eleven yes, 10 no. And that will be in the grey zone.

MS. BAL: Yeah. So, as mentioned earlier, with the grey zone, we'll put this out to comment and then you'll reconsider this measure during your post-comment call.

CO-CHAIR LANG: This is open for public comment.

MS. BAL: Operator, could you please

open up public comment?

OPERATOR: Yes, ma'am. At this time, if you would like to make a comment, please press star then the number 1. Okay. And you do have a public comment from Elizabeth Herman.

DR. HERMAN: Yes. Should I limit it to this measure or can I also comment on the previous measure?

MS. BAL: You can comment on any of the measures that have been discussed so far.

DR. HERMAN: Great. So, my name is Elizabeth Herman. I'm a senior scientist with CDC's National Asthma Control Program. And these are my opinions and do not necessarily represent official CDC statements. On measure 2852, optimal asthma care, I'd like to note that this is a patient-centered and measures a proximal outcome of care, asthma control, which fills an important gap in the current set of asthma quality measures. It's also well-grounded in the evidence and aligns with the National Asthma Education and Prevention Program Guidelines. It

also has been shown to motivate providers to incorporate an important, but underutilized tool for asthma care, the assessment of asthma control using validated instruments. This measure is analogous to and as important as the measure of blood pressure for the control of hypertension.

Regarding measure 2794, this measure offers two significant improvements over existing measures of emergency department use for asthma. One, it focuses on children with previous documentation of asthma and, thus, children who should be integrated into a system of care. And it addresses issues of gaps in enrollment and change in health plans, or churning, which are of concern to providers and health plans. Thank you for the opportunity to comment.

OPERATOR: Okay. You have another comment from the line of John Schall.

MR. SCHALL: Hi, thank you. I just wanted to make a somewhat more general comment to urge that we're mindful of some material and systematic technical discontinuities caused by

the administrative data coding. The definition changes that occurred October 1 of last year, when we moved from ICD-9 to ICD-10, and essentially there was one code that was fairly common, it's the chronic obstructive asthma that in ICD-9, 493.2X was considered an asthma code, but under ICD-10, it's considered a COPD code.

And so, as long as there's consistency between time periods and baseline versus performance, there's not a problem, but if you have a baseline using ICD-9 data and a performance period using ICD-10 data, then there could be an issue unless there's some type of consistency adjustment. And, you can adjust by -- there's some measures that combine asthma and COPD, in which case it's not an issue. If you count the particular code range as if it's COPD even though it's an asthma numeric code, it'll be in both sides and would be consistent. If it's not consistent, asthma is defined by any of the 493 codes. It may still be okay if it's not materially different.

So, for example, that code really represents both asthma and COPD and there's far less of that in children so it probably is not an issue on the pediatric measures, but could be elsewhere. The problem is, it may be a 20 percent shift in rates for asthma or COPD depending on whether it's consistent or not. And that's a concern for perhaps a couple of the measures that you're looking at over the next two days and for the future, because this is a standing committee, we're likely to see broad based measures like morbidity rankings for COPD rise because what used to be counted as asthma now is coming in as COPD.

It may also be an issue for other measures indirectly because COPD and asthma are currently both considered chronic conditions and the risk adjustment in a number of measures across various specialties include a count of comorbidities. And someone that has both asthma and COPD may have in ICD-9 two comorbidities and in ICD-10 only one for the same patient, just

because of the coding difference. And being mindful of that will be helpful going forward. Thank you.

OPERATOR: Okay. And at this time, there are no comments.

CO-CHAIR LANG: Okay. Thank you for the public comment. Rather than, we have our next measure 2816, but rather than placing that measure in between us and lunch, what we'd like to do is to take a ten minute break and each of us can grab our lunch and bring it back and we can have a working lunch while we're addressing that measure. Thank you. So, we're going to take a ten minute break and we get to extend our time with you, Larry. Thank you.

(Whereupon, the above-entitled matter went off the record at 12:32 p.m. and resumed at 12:44 a.m.)

CO-CHAIR LANG: We're going to start in a minute if everyone could please take their seats. The next measure, the next morsel for us is the measure 2816 entitled Appropriateness of

Emergency Department Visits for Children and Adolescents with Identifiable Asthma.

This is categorized as a process measure. Larry, would you like to go over this for two minutes for us please?

DR. KLEINMAN: Thank you. The first thing I want to say is we thought of this as a process measure because that's how I was trained to think. But it was raised by the committee that is this better thought of as an intermediate outcome measure.

And I would be prepared to accept that judgment if that were the case. I think there are, they are both there. I think this is an interesting measure. There are a paucity of appropriateness measures out there.

So it's been perceived as a need. This is again a Pediatric Quality Measures Program measure requested by, we were requested to look at overuse in ED asthma by AHRQ and CMS.

And it's a challenging thing to do. This is not a perfect measure. I think it's a

very good measure. It is, it's also important, there was a question about gaming that came up, the notion of gaming and false documenting or forcing documentation.

Yeah. Okay. Happy to do that later then. That's fine. Thank you very much. And this was developed for through same process and also endorsed by the expert panel.

CO-CHAIR LANG: Thank you. Dale, take it away.

CO-CHAIR BRATZLER: All right. So I strongly agree also that this is not a process of care measure. I embrace the argument that this is an outcome measure that focus on the appropriateness of emergency department visits for children and adolescents.

The denominator for the metric were those patients with an emergency department visit for asthma as a first or second diagnosis who meet other specified criteria for having identifiable asthma. The numerator for the measure is appropriateness as defined by one of

eight consensus appropriate use criteria that the expert panel defined.

They include hospitalization directly from the ED, documented physical exam findings consisting of respiratory distress including labored breathing retractions, accessory muscle use or markedly decreased breath sounds, O2 saturation less than 90 percent and ABG obtained, a consultation ordered and obtained with a pulmonary asthma specialist and order for an arterial blood gas or consult with pulmonary asthma, patient caregiver referral to an ED after evaluation from the PCP, patient or caregiver report of administering two or more doses of an inhaled rescue medication or parent or caregiver report that the child was in a pre-defined red zone on their peak flow measurement as a part of asthma.

So that's the numerator, denominator are again are those patients diagnosed, seen in the emergency department for asthma with that in the first or second diagnosis. Because this, I

believe, is an outcome measure it's an appropriateness measure, then the question on evidence becomes simply are there processes, structures, changes in care that could potentially impact the outcome for the measure which I believe there are. So that's my evidence discussion.

CO-CHAIR LANG: Are there additional, or are there comments, questions, yes, please.

DR. DIGIOVINE: So I'm understanding this measure as sort of we're asking whether patients are going to the emergency department when they should not be going to the emergency department based on their level of severity. What is the evidence that educating patients or giving them action plans or any of those things prevents those kinds of ED visits?

DR. KLEINMAN: The way that question was, just to be absolutely explicit, was posed to the expert panel was is the emergency department an appropriate level of care for a patient blah, blah, blah. So just not to put, to overstate it

in one way or another.

There is evidence that levels of appropriateness of a variety of things have changed, are changeable. There is evidence that more robust primary care systems have fewer kids go to the emergency department and when they do they are sicker.

But I would say that this measure is best interpreted in combination with that rate measure because a high level of inappropriateness may suggest that asthma outcomes are better than expected or they were thought based simply on that, but that the efficiency and the capacity of primary care, capacity and/or skill of primary care to manage the patients is not sufficient.

A very low level suggests that the primary care system may be working well, but that most of those visits are actually because the kids are sick and that maybe it's not actually specifically managing the asthma as well. So this is a -- it's somewhat of a nuance measure and I think it works better with the other one to

tell a story.

CO-CHAIR BRATZLER: Yes. So again, I think we were okay that a structural issue, if you had better access to primary care you might have fewer admissions to the emergency department. Access has been linked to rates.

So when we talked about this before that was our feeling that as an outcome measure the question is are there structures, processes, interventions that could impact the outcome. Our answer to that was, yes.

DR. DIGIOVINE: So let me start maybe asking a science question. So I have no doubt that there's lots of research on access to care decreases ER visits. Is there lots of science that says access to care decreases inappropriate ER visits where they actually look at down at this kind of detail was this an appropriate ER visit?

DR. KLEINMAN: I can answer that question. There is evidence limited that kids, that access impacts the severity of care with

which kids present to the emergency room in the way you would expect with this.

This particular measure and these details have not been studied. This is a new measure. It's a new formulation that came from the expert panel. Each of these was approved by the expert panel.

DR. GLOMB: Thank you. And following up on that this, I mean I loved everything about this. I intellectually agree with this approach. I think it sounds like a good measure.

But that last statement that you made is, at least my primary concern that I want to extrapolate in that, you know, we have three years ago we killed some measures because one facet of the components just you couldn't draw a bold straight line from Point A to Point Z.

And here I am concerned that this list of eight things, most of which I agree with, is arbitrary. It might be an indicator of appropriateness or it might not give a larger, you know, a step back.

And if we can't prove that all eight or seven of the eight or five of the eight are not valid measures of appropriateness can we then hold someone to the measure?

DR. KLEINMAN: I appreciate your concern. What I would say is this was designed to be an index at a population level, not to grade the care of any individual.

Again, these are system-level measures not even a hospital level measure. And as such I think the signal-to-noise is high and the validity comes from the fact that this was proceeded, that the expert panel was proceeded by a literature review, interviews with front line patients --

MS. BAL: We're on evidence. Please focus on evidence.

DR. KLEINMAN: Okay.

CO-CHAIR LANG: I have one more question for you, Larry, on number, I need to ask you this. Number five, consultation are obtained with a pulmonologist asthma specialist. Isn't an

allergist, immunologist a pulmonologist asthma specialist?

DR. KLEINMAN: I would consider you would be an asthma specialist, yes.

CO-CHAIR LANG: With that thank you, with that clarification we will proceed to vote on the evidence. According to the NQF algorithm for evidence this is eligible for insufficient with exception. So we must achieve more than 60 percent insufficient on an initial vote on evidence.

MS. BAL: So just a little additional clarity, if you are to vote and you vote low then the measure does not move forward. But if you vote insufficient then we would vote insufficient if you would like to pass it with insufficient with exception. Could you put your mic on?

CO-CHAIR BRATZLER: I need a point of clarity. If this is, if we're going to call this an outcome measure where does that show up as insufficient?

MS. BAL: So actually it wouldn't be

considered an outcome measure because we have to review the measure as it was presented. And since the developer presented it as a process measure we do have to vote on it as a process measure.

And if we don't move this measure forward there is opportunity for the developer to bring it back and make alterations as necessary. But since they did not provide us the evidence based off of it being an outcome there's no way for us to really evaluate it in that manner.

MS. GORHAM: And just a reminder, you have the algorithm at your seat if you would like to take a look at that.

MS. AMIRAULT: Okay. So polling will be open for the first evidence part for Measure 2816. Again --

DR. GROSSBART: Can you repeat what you said at the start of how, what each vote would mean to the measures, the result of the measure? Is that possible? I couldn't follow it.

MS. GORHAM: Could your question please, can you speak into the mic?

DR. GROSSBART: Can you repeat what you had said at the, as you introduced the voting?

MS. BAL: Yes. So if we vote high or moderate you are passing this measure on moderate saying that this has enough evidence to move this process measure forward.

If you vote low you're saying that the evidence provided is not applicable or it's not enough evidence or it was just insufficient indicates that you feel that there is not, the, basically there is not enough information to move forward, that whatever has been provided to you is not sufficient enough.

And so that, only if we have 60 percent or more on just insufficient shall we move on to an additional vote for insufficient with exception. However, if you wish to not go towards that you would vote low. Based on our algorithm it would not be eligible for high or

moderate.

MS. AMIRAULT: So we go ahead. For evidence for 2816 one high, two moderate, three low or four insufficient. Again, evidence for 2816.

(Voting.)

MS. AMIRAULT: Okay. Zero high, five moderate, seven low and eight insufficient.

MS. BAL: So since we did not get 60 percent on, sorry, insignificant we actually, this measure would go down because we would need at least 60 percent or more for it to move forward for insufficient with exception. So this measure does end with evidence.

DR. KLEINMAN: Is there an opportunity potentially for the Committee to revote because those who provided moderate might like, obviously there's more than 60 percent who wanted it to move on and if they move to insufficient that would allow the measure to move on?

MS. BAL: I will leave it up to the Co-Chairs on the Committee if they would like to

revote.

CO-CHAIR BRATZLER: I can only speak for what happened at our pre-meeting which was we didn't have this conversation that we had to evaluate this based on it being a process measure. Our feeling as a Subcommittee meeting beforehand was that this was clearly an outcome measure.

So we never had any of this conversation at all. That's what surprised me a bit here. We had the, our conversation based on this is an outcome measure, not a process measure. So I'm not trying to make the decision one way or the other. I just was a bit surprised by this.

CO-CHAIR LANG: I think if there's some misunderstanding regarding, I would say if there some misunderstanding regarding those who vote moderate we should, you know, have clarity. So I would agree that we can revote.

MS. BAL: Okay. So we're going to go ahead and revote on this evidence for this

measure. Again, with the algorithm your options really are just low or insufficient.

If we do not receive at least 60 percent insufficient votes than the measure will go down at this point. So give us one second to reset the voting and then we'll let you know when it's ready.

MS. AMIRAULT: You can go ahead and put your votes in, in the meantime. Sorry for the delay. Give us just one more second.

MS. BAL: Still having some technical difficulty. For everybody on the phone, we'll be starting in just a second with the vote.

CO-CHAIR LANG: Sorry for the delay. The battleground state of Ohio it's voting day today. I just got a, sorry, I just got a text from my wife saying that we just got a call from Arnold Schwarzenegger asking us to vote for John Kasich.

MS. BAL: With that we are now ready to vote. So please, as mentioned, the options are three low, four insufficient. If we have 60

percent or more for low than we will not continue discussing this measure.

But if we get 60 percent or more for insufficient, I'm sorry, more than 60 percent for insufficient we will do a secondary vote.

DR. DORMAN: Did Arnold have recommendation for us?

(Voting.)

MS. AMIRAULT: So zero high, two moderate, nine low and nine insufficient. And so we're at 45 percent insufficient.

MS. BAL: So with that on our second vote this measure does still go down and we can move on to the next measure.

CO-CHAIR LANG: Thank you. So at this point prior to moving on to the next measure we have the opportunity to discuss related measures, to discuss the three asthma measures that we've addressed this morning and into this afternoon including the Minnesota Community Measure, Optimal Asthma Control and the two measures pertaining to emergency department visit use and

appropriateness of emergency department visit use.

We're discussing this in terms of harmonization, yes?

MS. BAL: Yes, I just wanted to provide some clarification.

CO-CHAIR LANG: Please.

MS. BAL: That the 2816 since it didn't pass we wouldn't review it for related and competing. So at this point we're only looking at 2852 and 2794 as related measures. So if you could just bring up the decision logic.

So we had some questions earlier about what is the logic behind creating the related and competing. So this is what NQF considers when we make that determination comparing all the different measures. So first we really begin to look at does the measure address the same target population or the same measure focus.

If the answer is, no, then we consider the measures not related or competing. If the answer is, yes, we move on to the next aspect

which is does measure address both the same target population and the same measure focus?

If the answer is, no, we would start considering this measure for just related. And if the answer is, yes, then we would start considering it for competing. Then moving from there for the related measures I'll focus on that since we do not have any competing measures.

We would address with, if either of the same target population or the same measure and if they do then we would consider them related. And if they don't then we would not consider them related.

So I hope that brings a little more clarity about the process that we took to determine what measures were related and which were not. So with that we can open us discussion. Yes, go ahead.

DR. O'BRIEN: Just a question. It appears that we're only considering new measures in this. Is there a reason we don't consider existing measures also as far as the

harmonization?

MS. BAL: So we do consider existing measures as part of the harmonization. At this point when staff reviewed all the measures we felt that only these three were related.

However, if you think that there is an endorsed measure or a measure in our portfolio that would also be related or competing you are free to mention that now and we can open up the discussion to the Committee.

DR. O'BRIEN: So I think the optimal asthma control, one of those measures related to admissions and there's a PQI measure that is admissions for asthma.

MS. BAL: Could you provide the number please?

DR. O'BRIEN: Sure. It's 0283 and then 0728 are both asthma admission rates.

MS. BAL: Okay. We can definitely add that to the list of discussion points if the Committee as a whole agrees. So just to confirm you said 0283 and 0275?

DR. O'BRIEN: 0728 was the second one, asthma admission rates. They are PQI 15 and PDI 14. And we've got those tomorrow as a discussion related to competing measures to other, that we're discussing tomorrow.

MS. BAL: So since those measure those developers are, were not notified that the discussion would happen we should focus discussion on measures and we can do that discussion during the meeting tomorrow. Would that be okay?

Okay, perfect. Thank you. So then let's focus the discussion on the two measures in front of us which is 2852 and 2794. What we're asking of the Committee is do you feel that these measures need to be further harmonized or are they efficient the way that they are?

Is the developer, Jasmine from Minnesota still on the line?

CO-CHAIR BRATZLER: So I'll just make a comment that the data source for these two separate are very, very different. One focused

exclusively on the medical record, one exclusively based on claims analysis.

So the, their use would be very different. I mean Minnesota made it very clear that they did not have a unified, all paired, discharged data set to actually collect this data from so they relied on medical record review for 2852 versus 2794 which is based entirely on claims.

MS. BAL: Okay. So based on that I would say that you're saying that you don't feel that these need to be harmonized further, correct?

CO-CHAIR BRATZLER: I think it would be difficult to harmonize them further.

DR. GLOMB: The person who called in from, was it CDC also talked about how the Minnesota method was a patient-centered measure. So their claims base would be potentially outside that. So again, a little bit of a skewed look at the data.

DR. O'BRIEN: Looking at this

framework where is the data source? Which step is the data source?

MS. BAL: Okay, so we're getting the feeling that everyone is fine with these measures as is and we can move forward.

DR. O'BRIEN: Well I was actually asking for a walk through this step by step because I don't see the data source as one of the steps. These look to me, I mean they're addressing the same target population, the kids with asthma.

CO-CHAIR BRATZLER: The same target population although I know this issue has come up before for a whole host of measures where the condition or the patient population, so take nursing home versus hospitalized patients or others where the population of patients may be very much the same but the source of care or the site of care or the source of the data that actually used to populate the measures is very different. So I don't have strong feelings about it.

DR. NISHIMI: The, whatever you want to call it, it's not really an algorithm but the guidance does not, as you indicate focus on data source. That is one of the reasons that the Committee can cite as to why they feel that two measures are appropriate.

Different data source and then different foci. One was more patient directed as Dr. Glomb indicated and, you know, one is at the claims integrated healthcare system. One is focused on the providers, i.e. the front line docs. That's sort of what that other one was or groups.

The other is focused on integrated health systems and plans. So those are the reasons you can justify for differences. This helps staff sort the initial. But then your discussion is broader.

DR. DIGIOVINE: I guess I would just like to add my voice to saying I think these should be harmonized. I see them as relating to the same thing.

I don't think anyone brought up the idea that asking a patient whether or not they had an exacerbation was somehow a more valid measure than claims data. It was just brought up that it was probably as good or it might be as good.

And so I have my concerns that and I think the comments from the CDC were really about using an asthma control questionnaire which I don't think anyone disagreed with. I think that all of the disagreement was around the emergency room visit measurement.

And I think these two measures would work better if the first one stuck only to assessing control and we used the latter to measure ER visits.

DR. GLOMB: I second that. I really, that was my big problem with Minnesota as it sat today was it's just putting so much credence in the memory of the patient.

DR. NISHIMI: Any other comments?

DR. KLEINMAN: First of all I

appreciate the confidence in our rate measure. I just would note that there are different age populations. We go up to age 21. They go up to age 50.

And we do have that prior requirement which is in terms of identifiable asthma which I don't think is a problem. But it would well, it could conceivably and likely be defined differently at an older age.

DR. GROSSBART: I had a comment on the patient-centeredness discussion. The measurement developer didn't bring that up. So they weren't, and I don't believe that, they didn't bring that up.

And I know we worked with the Minnesota group and Cincinnati on aligning forces for quality. And the big appeal and pitch was that providers had more confidence in their own data than the insurance companies data.

And that's how the debate was. So I don't know that, you know, this is using the claims is patient-centered or using the ED self-

reported makes this more patient-centered. It's, and I don't know if that's what this comment --

DR. NISHIMI: That's fair. It's in the NQF rubric a patient-reported outcome is viewed so, but that's absolutely a fair comment. You're right. The developer themselves did not raise that.

MS. BAL: Okay. Were there any other comments? As always we recommend that the developer take back this feedback and when they do, you know, once you get this all processed if you bring it back to take it into consideration.

DR. NISHIMI: So what unfolds is that the developers are given this feedback. We ask them to have discussions and then report back as to why they feel they should or should not, you know, harmonize these.

The developers are under no obligation, obviously to do so. These are the Committee's recommendations. But they should be taken to heart because the measures have to come back for measure maintenance.

CO-CHAIR BRATZLER: So I just want to summarize that I think what I've heard though is fairly consistently for the Minnesota measure there is still concern about the self-reported ED inpatient and that if there's any way they can incorporate the claims-based approach to identifying those that would be recommended or certainly harmonization with what you've done with the pediatric population.

DR. KLEINMAN: We're not being asked to do anything at this point from our end. Is that, thank you, okay.

CO-CHAIR LANG: At this point I will pass the baton to Dale to proceed with additional review of measures pertaining to asthma. Thank you.

CO-CHAIR BRATZLER: All right. Thanks, David. So we're going to move ahead with Measure Number 0047, asthma, I'm sorry.

MS. GORHAM: Before we actually do that let us go over the differences between the maintenance measures and the new measures because

now we're about to begin maintenance measure review.

Okay. So now that we have our slides up, so as you know NQF reviews maintenance measures every three years or so. And this is to ensure that the endorsed measures reflect the current science and are reliable and valid and still meet NQF's criterion.

We do have a new maintenance process. And that process began, was effective October 1, 2015. So of course this project is included in the new maintenance process.

The maintenance measures use the same criteria as the new measures. But the emphasis is slightly different. So as you see on your screen for evidence there is a decreased emphasis for maintenance measures.

So if the developers have attested that the evidence is the same and you agree that there is no new evidence then we can say that we do not have to read this in the conversation or the vote for that. However, for a gap there is

an increased emphasis for maintenance measures and so we definitely want to look at gap.

So for scientific acceptability there is no difference for specifications. We still want to look at specifications for maintenance measures. However, for reliability and validity testing again if that is the same and you all, the developers have attested that is the same we do not have to revisit that conversation or vote for that.

However, we do want to address the questions of SDS for our trial period. So again, feasibility there is no difference. For usability and use there is an increased emphasis because the maintenance measures have been in use and so we do want to look at that data and there is increased emphasis on that.

Okay. And so now, Dale, we can move to our first maintenance measure.

CO-CHAIR BRATZLER: All right, very good. Thank you. So our first measure is 0047, Asthma: Pharmacologic Therapy for Persistent

Asthma from the American Academy of Asthma, Allergy and Immunology. And, Rebecca, I'll let you give it a brief overview of the measure.

MS. SWAIN-ENG: Thank you very much. So I'm here on behalf of the American Academy of Asthma, Allergy, and Immunology or AAAAI. Just to give you a little background the AAAAI just recently overtook this measure from the American or excuse me the American Medical Association's Physician Consortium for Performance Improvement in late 2015.

So AAAAI took over stewardship just the end of last year and is kind of new to this process. So we are just kind of getting our feet wet on the ground really learning what does NQF endorsement mean? How do we maintain endorsement?

What do we need to do to do a little bit of testing to make sure that we're up to date with the new NQF endorsement process? Since as Shaconna just mentioned it's really been updated here recently just at the end of last year.

So we're really happy to be able to present this measure for your review today for maintenance. This measure is currently part of the 2016 CMS PQRS program. It has been used in the PQRS program for the past several years.

It is one of the only asthma measures in the asthma measures group reporting option and is one of only two asthma quality measures in the program. It has been used in several AAAAI programs for quality improvement, maintenance of certification and as an AAAAI Qualified Clinical Data Registry or QCDR.

AAAAI is also aware that several other organizations have incorporated this measure into their internal quality improvement program. So it's definitely in use, which is one of the big criteria that you'll be reviewing today.

This measure has been previously endorsed by the National Quality Forum most recently in 2014. And this measure was recently updated. The specifications were updated. So the denominator was updated to take out the upper

limit of the age range.

So previously it had an upper age range limitation of 64 years old. So now it's just five years and older. And this was done to be able to be used in the PQRS program to include the Medicare population.

And also the numerator was recently updated just to be able to include more generic drug names and to better reflect what's going on in practice. So we're just really looking forward to hearing your input today and if you do have any questions please let me know.

CO-CHAIR BRATZLER: Okay. Thank you, Rebecca. So our two discussants are Curtis and Steve who is, take the lead.

DR. GROSSBART: I'll just, you just want to flip-flop back and forth?

DR. COLLINS: Well, I had that it's Christine and Crystal on the, it changed.

MS. BAL: That was a typo.

DR. COLLINS: Will you take the lead please?

DR. GROSSBART: Easy enough. So starting with the evidence the measure developer reports that there's no evidence changes and no evidence to call into question the measure.

And just, this is just to reiterate this is a process measure and the developer already made some nice comments on what minor changes have been made. And the evidence still seems valid.

CO-CHAIR BRATZLER: Okay. So any questions or comments about evidence?

DR. NISHIMI: So can I see a show of hands if people feel that we should revote and further discuss this, if you feel we should do that? Okay.

(Off microphone comment.)

DR. NISHIMI: Well based on what Steve has indicated and what you find in the --

DR. O'BRIEN: Do you mean revote on just this item, just Item 1A?

DR. NISHIMI: Just this item. We're just talking about evidence and the summary

that's in the preliminary analysis and what Steven just summarized indicates that the evidence hasn't changed. Developer attests it hasn't changed.

CO-CHAIR BRATZLER: So I guess the question, raise your hand if you think we need to revote evidence?

DR. NISHIMI: We're just talking about evidence.

CO-CHAIR BRATZLER: Seeing none we'll move on.

DR. NISHIMI: Right. So we'll move on to performance gap.

DR. GROSSBART: Performance gap. The developer shows that there is a gap and the Work Group discussion focused a lot on whether the measure has approached being topped off. So the developer's standpoint is that there is opportunity to continue to close the gap.

There's variation. There's still, as in the data presented, there's still variation particularly race, ethnicity and socioeconomic

status continue to have a disparate care. And so again, there still is a gap.

This measure probably in the near future will top out if the progress continues. But I know it was my feeling that there was still opportunity given, you know, given a still significant band between disparate groups.

CO-CHAIR BRATZLER: So, Bruno, did you have a comment?

DR. DIGIOVINE: I'm sorry.

CO-CHAIR BRATZLER: Okay. So, Steve, I just wanted to clarify. Is the total performance metric reported 99.3 percent? I saw that number quoted.

DR. GROSSBART: I would have to get into the data.

DR. COLLINS: Yes, that would be a question I would have for the developer as well. They have 88 percent for one measure, 71 and then a combined measure of 99.3, which that is getting very close, in my opinion.

And I expressed this in the Work Group

as well as being topped out or close to that discussion point.

MS. SWAIN-ENG: So that's just from one of the recent testing that we did just looking at one sample looking at one clinical, one large health group. And that's right.

When we looked at the combineds. We looked at one thing I didn't mention in my introduction. We looked at three different rates and that was the combined for any of the long-term control medications was the 99.3 percent within all of those different clinics.

However, we haven't had a chance to really get a lot of data from the PQRS program yet. So we just updated the metrics in 2014 to go up to the 50 year old age range and then in 2015 it included up to the 64 year old age range and now into 2016 there is no upper age range.

So we expect to still see a lot of variation. And so we think there is still going to be a lot of gap that is going to come from that. And one thing that we did note, I don't

know if you all had a chance to see this, is that when we looked at the data even from 2010 to 2013, the PQRS data from those age ranges only 38.7 percent of those physicians had at least 90 percent of a, a 90 percent performance rate or higher.

So there still is a significant portion of physicians that are, at least according to that PQRS data that aren't at that really high performance rate. So there still is room for improvement with this measure, we believe.

CO-CHAIR BRATZLER: So I guess has there ever been any discussion because, you know, I've actually looked at this metric before and there's a long list of controller medications. It's not that hard to pass this performance measure with the long list of, you know.

But when you get down to inhaled corticosteroids then the rates drop off. So has there been any conversation with the Work Group about whether, you know, you need or whether the

focus should be on some type of corticosteroids in these patients or inhaled versus just the long, it's a very long list of medications that you could pass this measure with?

MS. SWAIN-ENG: I wasn't privy to those specific conversations. That's something I could ask the Work Group and kind of come back to you all with that.

I know they wanted to try to be more inclusive to be able to meet the different individual patients that would be needing different medications based on comorbid conditions or patient preferences or, you know, what does their insurer cover to be able to meet the broadest range of medications that would be applicable for this patient population. That's a good question.

DR. JIMENEZ: Thanks. Would this be something that we would, should see in utilization or utilization on this one? I mean how is the measure being utilized right now?

MS. SWAIN-ENG: Do you want me to

answer that now or wait to --

DR. NISHIMI: Under use and usability.

DR. GLOMB: To follow on what you're saying, Dale, I wonder if we couldn't make a recommendation at some point that perhaps there needs to be a couple of numerators with this particular measure. For instance, all controllers versus ICS or whatever option it might be.

DR. NISHIMI: Sure, that's something that we can include in the report.

CO-CHAIR BRATZLER: Other comments? Okay. So I think we're ready to vote on gap.

MS. AMIRAULT: Okay. So we'll vote on gap for Measure 0047 and the options are one high, two moderate, three low and four insufficient. Again, for gap for Measure 0047.

(Voting.)

MS. AMIRAULT: Okay. Four high, 16 moderate, zero low and zero insufficient. And based on the percentage we can move along.

CO-CHAIR BRATZLER: Okay. We'll move

on to scientific acceptable.

DR. COLLINS: So for reliability testing the developers raised the age limit above 65. They tested for reliability with their beta-binomial analysis and reported reliability rates greater than 0.97 in all three categories.

The Work Group did not have any concerns with the reliability or the updated testing.

CO-CHAIR BRATZLER: Anything else, Steve? So any other comments about reliability? Okay. I think we're ready to vote.

MS. GORHAM: Before we continue with the vote just for record purposes I want to state that David Lang, one of the Co-Chairs is recusing himself from the discussion and the vote of this measure.

MS. AMIRAULT: Okay. So for reliability for Measure 0047 the options are one high, two moderate, three low and four insufficient. Again, reliability for 0047.

(Voting.)

MS. BAL: Ella, if you could please vote. Sorry, never mind. We've received.

MS. AMIRAULT: Okay, 12 high, 8 moderate, zero low and zero insufficient. You can move along.

CO-CHAIR BRATZLER: Okay, to validity.

DR. GROSSBART: So the, excuse me, validity testing phase exclusions appear appropriate. Risk adjustments not applicable. The biggest, the validity of the measure was, I think Curtis already touched on this, the beta-binomial analysis some concern of the strength of that relationship.

Developer argues that .7 indicates sufficient reliability to distinguish among physicians. So there's, so those are the main points that I recall from the conversation, Curtis.

DR. COLLINS: Yes, the Work Group really didn't have, discussed reliability. Validity was the same way. I think that the developers, we thought that they had done

appropriate testing on the updated age groups methodology. And we really didn't have concerns with validity either.

DR. NISHIMI: I just want to remind the Committee that they did data element level validity testing not score level and face validity. So the highest eligible vote, based on the NQF algorithm would be moderate. So your options are moderate, low and insufficient.

MS. AMIRAULT: Okay, so for validity for Measure 0047 the options are two moderate, three low and four insufficient. Again, for validity for 0047.

(Voting.)

MS. AMIRAULT: Okay, so zero high, 17 moderate, three low and zero insufficient. And based on that percentage we can move along.

CO-CHAIR BRATZLER: Okay, feasibility.

DR. COLLINS: From a feasibility standpoint it's electronic data currently being captured in, you know, quite a few sources. We did not have concerns with feasibility. Pretty

straightforward. I don't know if anything has changed since the last review.

CO-CHAIR BRATZLER: I would assume this is a metric that can be reported via codes also.

MS. SWAIN-ENG: Yes.

CO-CHAIR BRATZLER: Some physicians may report this using claims, not using electronic health record data. Some physicians may report this metric as a PQRS measure using their claims data rather than using electronic health record data to actually report it.

So any other questions or comments about feasibility? All right. We can vote.

MS. AMIRAULT: Okay. So feasibility for Measure 0047. The options are one high, two moderate, three low and four insufficient. Again, feasibility for 0047.

(Voting.)

MS. AMIRAULT: Okay, 17 high, 3 moderate zero low and zero insufficient. And based on the percentage we can move along.

CO-CHAIR BRATZLER: Okay, usability and use.

DR. GROSSBART: In the area of usability and use this is already a public reported measure, part of the PQRS measure set. It's been used in some payment programs, public reporting and the Work Group did not see any unintended consequences of continued use.

DR. COLLINS: I'll add too that there was improvement in performance from 2011 to 2013. There was nearly a 20 percent improvement in performance on this measure reported by the developer. So it seems to be usable.

CO-CHAIR BRATZLER: So I'm curious what data source is used to demonstrate that improvement?

MS. SWAIN-ENG: I think the one he is referring to is the Maintenance and Certification Program through ABMS. So using data from eCITY who is one of the vendors that works with a lot of the medical specialty societies to be able to report back which is a dashboard where you are

able to really see individualized physician reports. This is also seen in an aggregate report on your membership.

CO-CHAIR BRATZLER: And I make that point so that's very helpful just as you see with that program Maintenance and Certification and as you see with PQRS right now selection to the measures is left at the level of the individual practitioner or the group.

And so I will tell you as a group quality director that we pick the measures we look good on to report and we don't necessarily report the ones we don't look good on and I think that's probably consistent across most plans. So there's probably greater opportunity for improvement if everybody had to report the same metric than the pure voluntary reporting or it's not voluntary but you pick the measures because you get the option of choosing performance metrics that you choose to report.

Any other comments about use, usability? And if not we'll go ahead and vote.

MS. AMIRAULT: Okay, so usability and use for Measure 0047. The options are one high, two moderate, three low and four insufficient information. Again, usability and use 0047.

(Voting.)

MS. BAL: Susan and Ella, could you please send in your votes via the chat? We haven't received them yet. Thank you.

DR. KAZEROONI: I already sent mine. I'll send it again.

DR. POLLART: Yes, I did the same, but I just resent.

MS. BAL: Thank you.

MS. AMIRAULT: Okay, 15 high, five moderate, zero low and zero insufficient. And based on the percentage we can move along.

CO-CHAIR BRATZLER: And the last vote then would be just on overall suitability for endorsement. So is there any discussion?

MS. AMIRAULT: So for 0047 the overall --

CO-CHAIR BRATZLER: And I guess the

only thing we, I think we have the recommendation that the developer consider whether there should be two separate numerators for this particular metric based on the very high overall rate of the combined. So just at least as a consideration.

But any other conversation or comments back for the developer? Two or more, yes, right. Any other comments? Okay.

MS. AMIRAULT: So for the overall suitability for Measure 0047, one for yes and two for no.

(Voting.)

MS. AMIRAULT: Did anyone in the room not vote yet? Okay. So we have 20 yes and zero no for 100 percent.

CO-CHAIR BRATZLER: All right. Thank you, Rebecca.

MS. SWAIN-ENG: Thank you very much.

CO-CHAIR BRATZLER: All right. So the next measure that we're going to consider is Number 1800: Asthma Medication Ratio (AMR) from the National Committee for Quality Assurance.

And I'll repeat the announcement. Dr. Lang also is conflicted on this measure and will be recusing himself from the conversation or vote. So we have Lindsey Roth and Ben Hamlin as developers to present the measure.

MS. ROTH: If it's okay I'm actually going to provide an overview for the two asthma measures right now instead of separating those out. So we have two health plan level asthma medication measures.

And these improve upon an older measure of ours that had assessed whether patients were only dispensing controller medication at least once during the measurement year. But these two asthma measures that we will discuss today are higher bar than the older measure that we found had really high performance.

And they each focus on slightly different aspects of asthma care. So the first is the asthma medication ratio measure. And this looks at the percent of patients with persistent

asthma who have a ratio of controller medications to total asthma medications of .5 or higher.

The measure is an intermediate outcome measure that identifies people with persistent asthma who are well controlled and are not relying on rescue medications to control their symptoms but rather are following the recommended daily use of controller medications.

There are at least three studies that found that patients with a medication ratio above .5 are less likely to have a subsequent ED or hospitalization for an asthma exacerbation. And the summary wasn't included in your materials. But I did want to just point out that there are existing studies and we can provide more information about them if you would like.

The second measure is the medication management for people with asthma. And this measure looks at the percent of people with persistent asthma who were dispensed controller medications that they remained on throughout the measurement year.

So patients need to be on a controller medication in order to be in this measure. And the, what we're really assessing here is adherence to their controller medication to see who is really taking it as recommended.

And to assess adherence we used a proportion of days covered method which means calculating the percent of days that the patient was covered by at least one asthma controller medication during the treatment period. And we define the treatment period as the first day during the year that an asthma controller medication was dispensed through the end of the measurement year.

And this measure has two rates. One for the percent of patients who were on the asthma controller medication for at least 50 percent of their treatment period and the other rate is looking at those who were on it for at least 75 percent of their treatment period.

And I also just wanted to point out that we did provide a summary of studies that

have examined the link between the 75 percent and the 50 percent adherence rates in outcomes. And this was to show that there is some mixed evidence in this area.

So for example, a study came out last year. The primary researcher was Yoon. And they did not find a link between the 75 percent adherence rate and outcomes. They did however, find a link between the 50 percent adherence rate and ED visits.

But there have been several other studies that have validated the link between the 75 percent and 50 percent adherence rates and outcomes. So I just wanted to mention that we did have the opportunity to meet several times with the researchers on the Yoon study to discuss their analytical methods.

And they received our recommendations very well and they're actually conducting further analyses. So we're looking forward to seeing the results as well as any new emerging evidence that might come out in this area.

So just again, both measures are health plan level that use administrative and pharmacy data. And they are both stratified by four different age groups. And this is for several reasons.

The first is that we wanted to align with the asthma medication management guidelines which are separate for children and adults. And then the second reason is that the measures are used in programs specific to children and adults. And so for that reason we also wanted to see the rate separated by age.

CO-CHAIR BRATZLER: Thank you. So I have as the leads for the discussion Christine and Crystal.

DR. RILEY: Okay. I'll start with the evidence. The developer indicated that the updates to the evidence were that the 2007 guidelines cited updates for those that were published in 2004.

The guidelines being those created by the National Heart, Lung and Blood Institutes,

graded Category A where the guidelines reference a little over 550 studies that are related to pharmacologic therapy for asthma. So that seems to be the only change indicated for the evidence.

CO-CHAIR BRATZLER: So since this is a maintenance measure again we'll have that conversation about do we need to revote on the evidence and raise your hand if you think we need to revote on the evidence.

(Off microphone comment.)

CO-CHAIR BRATZLER: We're talking about 1800 only. We'll take the second measure. We're just going to talk about 1800 only for now. So anyone think we need to revote the evidence at this point? Okay, we'll move on then.

DR. SCHINDLER: Thanks, Lindsey. So in terms of the performance gap, as Lindsey stated this is a health plan level. So the only gap that really is identified is between the different types of products, so commercial product versus Medicaid, Medicare.

And certainly like we'll back up to

the raw numbers there are gaps that have been consistent throughout 2012, '13 and '14. There's not been a lot of movement on those. It's a pretty steady gap.

Unfortunately because it is a health plan level measure there really aren't additional data looking at demographic information. I think it would be really helpful if we could see something stratified by race, ethnicity, urban versus rural, age.

But it was really clear from them that this evidence is really hard or these data are really hard to extract given it's a plan data. It's just not reasonable to collect at that level. So disparity is macro but we don't really have the, you know, really specifics.

CO-CHAIR BRATZLER: Lindsey, any questions, comments? Anybody else have any questions? Yes, Bill.

DR. GLOMB: Just on the plan level if it's a Medicaid plan it's required that we have, that we collect that data. So at least a subset

of the plan data could be broken down demographically.

DR. DIGIOVINE: I guess to harmonize a little with the one we just talked about is there a reason that this ends at 65 and doesn't extend beyond 65?

MS. ROTH: So when we originally developed the measure we did make the decision to end at 65. And this was mainly due to some concerns we had heard about the older population and possibly, you know, misdiagnosis with COPD versus asthma.

And this measure is, you know, narrowly focused on asthma. So at the time we had made the decision to cap it although we are, we do have some data looking at the possibility of expanding it to not have an upper age limit and in the future we may decide to do that.

DR. DIGIOVINE: I just, 65 is a funny cut off. If that's the concern I would use 40. I think we typically use 40 if that's the concern.

MR. HAMLIN: So the original version of this measured population did actually cut off at 40 and several years ago we retested to look the kind of diagnosis between the asthma and COPD and found out that those disparities in the diagnosis really didn't exist.

And so we wanted to align the age stratification as much as possible to the product line that we report on as we mentioned before for the reporting programs. We do have an age strata from 50 to 64 I still believe in the measure that kind of addresses the potential differences in that population.

And that still is existing because we believe there may be some. But it hasn't born out right now. We do actually exclude anyone with COPD or any kind of chronic obstructive disease, so, yes.

CO-CHAIR BRATZLER: The only other question, I don't know if this is, it's not a gap. But, you know, the question does anybody ever graduate out of the measure or the

denominator?

So a patient has a diagnosis of asthma but, you know, gets better. I mean it doesn't become clinically relevant or they get where they move to get rid of their cap. They do something that's their triggers that, you know, don't require them to be on therapy. Does anybody ever graduate out?

MR. HAMLIN: Well according to our chair who is a pulmonologist generally they graduate into obstructive disease not generally, you know. So we do is, the number of exclusions for COPD does increase as the, as the age increases and it gets fairly significant I think over 65.

DR. DIGIOVINE: And I'm assuming the denominator is based on some set period of time with the diagnosis of asthma.

MR. HAMLIN: It's a two year denominator. So we work over multiple years.

CO-CHAIR BRATZLER: So any other comments, questions about gaps? Okay.

MS. AMIRAULT: Okay, so voting for performance for Measure 1800. The options are one high, two moderate, three low and four insufficient. Again, performance gap for Measure 1800.

(Voting.)

MS. AMIRAULT: Just to double-check, did everybody in the room vote? If you could just send again?

DR. NISHIMI: Please repress.

MS. AMIRAULT: Okay, six high, 14 moderate, zero low and zero insufficient. And based on the percentage we can move along.

CO-CHAIR BRATZLER: Okay, move to reliability.

DR. RILEY: Sure. So looking at reliability the developer noted that the specifications have not changed since the last endorsement submission. Reliability testing has been updated slightly to reflect new data from HEDIS from 2014.

But the developer also notes that the

reliability results still range from .93 to .97 indicating strong reliability.

CO-CHAIR BRATZLER: Any questions or comments about reliability? Seeing none, let's vote.

MS. AMIRAULT: Okay. So voting for reliability for Measure 1800. The options are one high, two moderate, three low and four insufficient. Again, reliability for 1800.

(Voting.)

MS. AMIRAULT: Okay. We have 15 high, five moderate, zero low and zero insufficient. And based on the percentage we'll move along.

CO-CHAIR BRATZLER: Okay. And so we'll move to validity.

DR. SCHINDLER: Probably the biggest threat to validity is the percentage of people excluded. So in the commercial plans they said about 25 percent of individuals are excluded while on Medicaid about 18 percent.

And interestingly it seems like a disproportion of those were in that older age

group of 51 to 64 versus the five to 50 year old population. I think it's important just to look at that, you know, up to a quarter of the patients are excluded. Otherwise there was no big change in the validity from the first go around.

CO-CHAIR BRATZLER: Lindsey, did you have any comments? Any other comments about validity, questions? Yes, sir.

DR. LAMPONE: Yes, I did have a question about the medication adherence ratio that you had used. And I think in Yoon there were some questions regarding when patients are prescribed either at the beginning of the period of end of the period and what that does to the measurements.

And in fact that study, that Yoon study had a lot of concern about that measure. Can you comment on that and what your interaction, what, you had mentioned that you have spoken to some of those investigators?

MS. ROTH: So actually the Yoon study

was more related to the measure that's coming up next, the adherence measure. So I mean I can comment on it now.

DR. LAMPONE: No, I got confused then on the measurement ratio that you were using.

MS. ROTH: On the ratio, yes, so it's a ratio of controller medications to total asthma medications, so controller and reliever.

DR. LAMPONE: And actually Yoon supported that as a much more accurate study across populations.

MS. ROTH: Right, yes. There have been a few studies that have found the validity of that.

CO-CHAIR BRATZLER: Other comments or questions about validity? Seeing none, let's vote.

MS. AMIRAULT: Okay. So for validity for Measure 1800 your options are one high, two moderate, three low and four insufficient. Again, validity for 1800.

(Voting.)

MS. AMIRAULT: Okay, eight high, 12 moderate, zero low and zero insufficient. And based on the percentage we'll move along.

DR. RILEY: For feasibility the developer notes that all of the data is generated during care processes and are currently included in defined fields in electronic claims. They also utilize independent audits to verify the HEDIS specifications are being met and that they do receive real time feedback from users of the measure.

So it doesn't seem like there are any concerns there with obtaining the data.

CO-CHAIR BRATZLER: So I'll raise the one issue that we talked about fairly extensively this morning and that's meds that may not show up in claims. It sounded like there are some periodic audits.

So I'm just curious, have you ever looked to see whether there's evidence patients were on meds that don't show up in your claims data files?

MR. HAMLIN: We've been increasingly interested actually for the next measure about those patients who don't meet the 50 percent adherence rate and finding out whether that's a data issue or a management issue. We haven't identified the perfect way to do that yet.

We're still looking. That's all I can really at this point. The consensus panel we used, the advisory panel is not convinced that the majority of asthma medications are, you know, sort of freebies or giveaways or low cost things. But that's just their highly respected opinion, professional opinion at this point.

CO-CHAIR BRATZLER: So other questions about feasibility? Okay.

MS. AMIRAULT: Okay. So voting for feasibility for Measure 1800. Your options are one high, two moderate, three low and four insufficient. Again, feasibility for Measure 1800.

(Voting.)

MS. AMIRAULT: Okay, 15 high, five

moderate, zero low and zero insufficient.

CO-CHAIR BRATZLER: Okay. To usability and use.

DR. SCHINDLER: So this is a maintenance measure there's a higher emphasis on this. And this current measure is being used widely and publicly report. It's being used at the health plan rating annual state of the health care quality health plan accreditation and quality compass.

The developers included lots of links, very easy to find, easy to identify. It's a really pretty straightforward process measure. There's been, they cited a little bit of movement in the Medicaid population not so much in the commercial.

But it's small. But certainly it seems like it's being widely used.

CO-CHAIR BRATZLER: I'm sorry. I just want to ask quickly, clarify something you said about not much movement in the metric, any concerns about why it's not moving at all?

DR. SCHINDLER: I don't really know why it's not moving. I think, I don't know that it's particularly problematic. I think the hope that would be by measuring it we're going to start addressing the gaps.

There still remains a very wide gap between commercial product and Medicaid, Medicare. And they cite that there's no movement but that the numbers, I didn't run a statistical analysis on it.

But they clinically didn't feel different, 57, 65 back to 59 percent.

CO-CHAIR BRATZLER: Bill, I'll go to you and then we'll ask the developers.

DR. GLOMB: With reflecting the Medicaid there's been a push in the Medicaid managed care programs to use this measure. And we're certainly using it as an indicator with, in quality care.

And now that payment models are beginning to evolve and preferred provider networks we've actually set this as one of our

measures to be a preferred provider of asthma care in our network. It's a very young measure and we've been rolling it out and people are starting to pick it up I think.

It is for the last couple of years been relatively stabilized. As it gains traction I think we'll see better improvement as it gets more traction, more programs.

CO-CHAIR BRATZLER: Any other questions or comments? Okay, let's vote on use.

MS. AMIRAULT: Okay. Voting for usability and use for Measure 1800. Your options are one high, two moderate, three low and four insufficient. Again, usability and use for Measure 1800.

(Voting.)

MS. AMIRAULT: Okay, for results we have 13 high, six moderate, one low and zero insufficient and based on the percentage can move along.

CO-CHAIR BRATZLER: Okay. And then overall rating, suitability. Any other comments

or questions, concerns or anything else to relay to the developer? All right. You guys are quiet. So we'll move on with the vote.

MS. AMIRAULT: Okay. So voting for overall suitability for Measure 1800. One for yes and two for no.

(Voting.)

MS. BAL: Ella, we have not received your vote. Please vote.

DR. KAZEROONI: I already did that one. I'll do that again.

MS. BAL: Thank you. We received it, thank you.

MS. AMIRAULT: Okay, 20 for yes and zero for no, making it 100 percent.

CO-CHAIR BRATZLER: All right, very good. Thank you. So we'll move on to the discussion for the next measure which is 1799. You've already given us a nice overview of the measure.

Anything to add to that? Okay. So our discussants are Thomas and Curtis and we'll

start off with evidence.

DR. COLLINS: So, you know, I think this is one that potentially we could explore again based on the Yoon study that was mentioned here previously. For those familiar with the measure they look at the numerator at the 50th and 75th percentile compliance measure.

This is a maintenance measure and it sounds like when it was approved before the level of evidence was questioned on those markers previous but was ultimately approved. So the Yoon study, correct me if I'm wrong, looked at this exact measure in over 30,000 patients.

And looked at whether the process related measure can be tied to really any sort of outcome. And the developer lists that at the 75th percent mark there was no difference in hospitalizations. At the 50 percent mark there was no difference in hospitalizations.

But there were fewer ED visits. That may or may not be clinically significant. The differences in ED visit rates was 3.7 and 4.2

percent. So just a half a percent difference.

And as far as what the authors of this very large paper say, they say there's no meaningful clinical outcome. If you read into it a little bit more some of my concerns are that this measure could actually increase costs and resource utilization and then potentially adverse events by actually increasing therapy or increasing medication use.

And the authors, while they don't provide a lot of great data for that do allude to that. So I think the level of evidence has certainly changed since the last review.

The developers had mentioned some follow-up or some additional studies, although I didn't see those in my review here. So we have a very large study and I think it certainly is noteworthy.

CO-CHAIR BRATZLER: Did you have anything different, Tom, or anything else?

DR. LAMPONE: I would agree with Curtis. Also the basic tenant of the study cited

that the measure doesn't take into account other triggers of asthma such as viral illness or other environmental issues that may trigger an event and does not take into account the treatment of mild asthma where daily controller use may not always be appropriate.

And that's where I think they came to some of the conclusions that this may drive utilization in some asthma populations that would be inappropriate.

CO-CHAIR BRATZLER: And so any other Committee Members have questions and then I'll ask the developers to respond to some of those comments. Go ahead, Lindsey.

MS. ROTH: Sure. So first I just wanted to point out that I believe it's on Page 3 of your measure worksheet is a summary of some other evidence looking at the links between the adherence rates and outcomes.

But going back to the Yoon study so one of the reasons that we are glad that we had the opportunity to meet with the researchers is

because we thought that there were a few limitations of the study. So the first one was that they were using an older version of the specification that had an error in how you calculate the denominator for the measure.

And it actually we thought was, could potentially be a, you know, a pretty big flaw because it was looking at how asthma or inhaler, canisters for inhaled medications were dispensed which is an important part of how you decide or calculate whether patients have persistent asthma versus the more, you know, mild intermittent asthma.

And the other, I think one of the other issues was they didn't control for disease severity which some of the other studies that did find a link to outcomes had done. And so they were also interested in going back and rerunning some data and looking at more recent data and addressing these two methods. So I think, do you want to add anything?

DR. DIGIOVINE: Can I ask, given

obviously there's a lot of questions about these cutoffs, why did you choose to make this a binary outcome instead of just saying it's a continuous measure of compliance and we would measure it like any outcome or any process we measure rather than a sort of yes/no?

MR. HAMLIN: So that's actually a great question because one of the things during the development process when we were looking at whether a medication adherence measure is appropriate for asthma as a condition, particularly persistent asthma, the group felt very strongly that just setting a single threshold like an 80 percent as you may see for diabetes adherence was really not appropriate.

And what they really wanted was a quality measure that essentially stratified the population into those who met a high level of adherence, those who sort of met the threshold for questionable whether it was adherent or not and those, the 50 and 75 percent were completely set as potential thresholds based on their own

judgments.

And we then have proceeded to look at the data as it comes in year over year. But again, it's a very young measure. So there's not a lot of studies that have been able to be completed in this short period of time.

The way I like to think about the measure is a 75 percent threshold tends to act as a performance rate and that's the one we use for the HEDIS accreditation scoring and the other programs where we're comparing plans. The 50 percent rate, and this is what we've been talking about with the Yoon researchers is actually it's those who are under 50 percent you probably want to spend more of your time focusing on whether it's because they're not taking their meds or they shouldn't be taking their meds.

And that's sort of the population of interest, if you will. And so that rate remains in the measure and is popular with some people and not with others. And so we wanted to leave the measure until we had some good data and some

good evidence to show that those rates were either inappropriate or should be adjusted up or down one way or the other. And that's basically where we are in full transparency.

DR. DIGIOVINE: Okay. Just to follow up just because I'm, this number, I understand you're measuring sort of days covered with a controller medication. Are you assuming that everyone has a prescription or how do you know this is, is it a compliance or is it a physician ordering issue?

MR. HAMLIN: It's based on dispensed events. So it's only if they pick it up. And we start the clock at the first dispensing event. So the date they pick it up, the first date they pick up their medication in the measurement year is when we start the count effectively.

It's admin claims so we have no way of understanding the orders unfortunately at this time.

DR. LAMPONE: So for clarification the adherence rate is calculated if a patient is

given an index prescription in October, October 1 they and they get that filled there's then three months in the treatment period and three months of active medication. And so they would be deemed compliant?

MR. HAMLIN: They would be. And we actually, we were very curious about that when we tested this measure with nine health plans when we first developed it. And we actually, we are sort of looking at when the index events occurs for most of the, majority of plan members and they actually happen the first quarter of the year.

So there are a few members who do meet the criteria. Remember we have a two year denominator for this measure. So you have to have two years of diagnosis or medications or combinations of events that get you as a persistent asthmatic.

And then so for that population really, I don't remember what the percentage? I don't remember the percentage numbers were. But

a very high percentage picked them. Their index prescription is the first quarter of the year, the first three months of the year tapering very rapidly into the second quarter and very few actually hit the fourth quarter.

DR. LAMPONE: So to be compliant in that example I used would the patient have to fill three prescriptions or if one was filled but the other two missed did you measure consistency or compliance with adherence?

MR. HAMLIN: So the measure requires a proportion of dates covered from the index prescription through the end of the measurement period. And we use the data supply field or the definitions that are in the measure because some of the data supply fields for canisters are a little wonky.

And so we use those calculations to determine the days covered, if you will, for each prescription. And many of them are not on any prescription. So in many cases for someone who is in the fourth quarter would get one

prescription and probably make it through the end of the year.

We don't count overlap. So if they get multiple prescriptions they're only allowed one controller per day, if you will, depending on which one that is in case of the medication switching issue and other things like that.

DR. LAMPONE: Okay. So I was having a little problem hearing you. So they have to have one prescription filled and it doesn't necessarily have to be for 30 or 90 days. You just count it across the board.

MR. HAMLIN: They just have to have one canister coverage per day of that period between the index prescription and the end of the measurement period.

DR. LAMPONE: Okay. So they would have to have, well three then, three canisters, okay.

DR. DIGIOVINE: Other comments or questions about evidence? So I think this one we want to revote because of the change. And I

assume, do you have any sense of your time line with Yoon in terms of the reanalysis of the data? I know it's not your data.

MS. ROTH: We're hoping soon. But, yes, we don't have a specific time line.

DR. DIGIOVINE: Okay. All right. So everybody agree I think we should revote the evidence on this particular metric. I'm seeing lots of heads nodding so I'm going to take that as a, yes.

MS. AMIRAULT: Okay. So voting for evidence for Measure 1799. Your options are one high, two moderate, three low and four insufficient. Again, evidence for 1799.

(Voting.)

MS. AMIRAULT: Okay. One high, 11 moderate, five low and three insufficient. And this will be a grey zone.

MS. GORHAM: Just for record purposes again David Lang has recused himself from discussion and vote of this measure as well.

DR. COLLINS: All right. So we'll

move on to, yes, performance gaps. As the developer has said on one of their previous measures, they have commercial data and Medicaid or Medicare, Medicaid data looking at differences and trying to account for performance gaps or opportunities for improvement.

It looks like between the commercial and the government payer there is, you know, differences both at 50 percent and 75 percent compliance ratios. And it looks like there's opportunities for improvement.

But there's no reports on disparities other than what's listed and what was listed before. And please correct me if I'm wrong.

DR. LAMPONE: I have nothing else to add to that.

CO-CHAIR BRATZLER: So any other questions or comments about gaps, performance gaps? Similar to the previous measure differences by plan but no other stratification data.

MS. AMIRAULT: Okay, voting for

performance gap for Measure 1799. Your options are one high, two moderate, three low and four insufficient. Again, performance gap for 1799.

(Voting.)

MS. AMIRAULT: We have five high, 12 moderate, three low and zero insufficient. And based on the percentage we can move along.

CO-CHAIR BRATZLER: Okay. We'll move along to reliability. Tom or Curtis, we're ready for reliability.

DR. LAMPONE: You ready to go? Okay. All right. Next section is reliability and I'll also after that discuss validity. And we touched upon this a little bit with the measure cut offs for the numerators at 50 and 75 percent of proportion of days covered.

There were no changes in this since the last endorsement. The denominator criteria had not changed. There's been some updates since the last submission but no big changes. I didn't have really anything else to add to this unless you do, Curtis.

DR. COLLINS: No, and the Work Group I don't believe did either.

CO-CHAIR BRATZLER: Any other questions or comments about reliability? Any other questions, comments? Okay. Go ahead and vote.

MS. AMIRAULT: Okay. For voting for reliability for Measure 1799. Your options are one high, two moderate, three low and four insufficient. Again, reliability for 1799.

(Voting.)

MS. AMIRAULT: Okay. Five high, 14 moderate, one low and zero insufficient. And based on the percentage we can move on.

CO-CHAIR BRATZLER: Okay, validity.

DR. LAMPONE: Okay. So moving on to validity. We touched upon this a little bit as well looking at the thresholds of the 50 percent and 75 percent.

There was note made that the field test examined ED visits for the populations above and below 50 percent and then there was two new

studies reported by the developer that looked at the proportion of days covered based on the index prescription.

When you look at the testing for this I was drawn to the, first to the Pearson correlation coefficients in which you look at the medication adherence and asthma medication ratio and the asthma medication ratio, you know, drops below the .3 percent threshold. And partly the other reason why that caught my eye is that the asthma medication ratio has been looked at in other studies as well.

And in the Yoon study they cite other particular studies that have looked at this. And particularly those studies address race and ethnicity through over a gamut of asthma patients and also had comorbidity index, the Charlson Comorbidity Index as well put it into that analysis.

So I had the sense that when you look at, at least the asthma medication ratio by itself it may be a more reflective measure of

actual compliance as opposed to the proportion of covered days.

And when you further look at that data where you have it somewhat risk adjusted and I know that wasn't part of the particular measure, it puts another level on information and outcome data, in this case exacerbations in ER visits or hospitalizations into a, I think a more tangible type of a perspective.

So I think just that nuance in that measurement kind of, in my mind puts a question to the validity. And also when you look at the outcome data, again as Curtis mentioned, there was no difference between the adherent and nonadherent groups as far as outcomes.

So you look at that and that kind of gives you another reason to pause and say is there something in the measure or is it something that is confounding the population and the added factors in there that is not taken into account.

CO-CHAIR BRATZLER: Curtis, did you have anything? So I'll turn it over to you in

just a moment. But other comments about validity?

So I mean I think this is a really important conversation about are we actually measuring something that will predict outcomes for the patients. And you would certainly think, I think if I heard you correctly, I was looking at it, I haven't looked at the risk adjustment.

Is there risk adjustment, any at all and so from a claims perspective things like ED and inpatient visits would be something relatively well available to most health plans as a predictor of risk. So do you have any comments?

MR. HAMLIN: Yes, so I mean it's not risk adjusted other than the fact that it's heavily stratified by both product line and age group which I call the poor man's risk adjustment for the health plan because really when you understand without going into the whole NCQA submission process that on the plan's report they're done on an enormous, usually a state base

because of the confines of the health plan contracts.

And so each health plan may have a submission for each state and that's the level of data in the aggregate we're getting. And so, you know, we have, we do risk it as other measures of resource use for asthma population.

But for the things like when we're looking at adherence we really try and focus in on the proxies, if you will, for the, you know, for the quality of care being delivered and there probably is some data noise in there and there probably are some other places to do.

But we, our current philosophy is that the plans are well positioned to further investigate as I mentioned on the 50 percent rate of the patients that are not. And without us sort dictating at the point of care what the physicians should be doing we kind of try and use these proxies, if you will.

So the asthma ratio is a proxy of asthma control using claims data that's the best

data we have for this kind of condition. And the same with this adherence. You know, I would love to have a much better measure using much more distributed data.

But in the admin claims this is really the best approximation we can get to as far as measuring the kind of care processes that we think lead to better asthma care. And I think there was an earlier comment about the last round and it was not a fan favorite.

But it was filling a huge gap and a need of people felt there weren't any other asthma measures out there that were doing anything similar. And so it was an important thing to just sort of, you know, get out there and they thought it was, met the level to be endorsed, if you will, because it was, there's really nothing else out there really that does the same thing.

It was an improvement over the measure that just said do you have asthma, do you have one medication, which is the prior measure that

had topped out basically for us. So we replaced it with this one setting the bar a little higher.

MS. ROTH: I also did want to add too that so for the Pearson correlations, so as you noted we looked at the measures are stratified by age but we did the correlations looking at the total rate between the medication ratio and medication adherence measure.

But after we had submitted the forms we did go back and look at the correlations between the two measures within each age strata. And what we actually have found was that the correlations were much higher for the children, the adolescents and the younger adults.

So the correlations were closer to .4 and .5 within those age groups. And it was really the older adults where the correlations were not as strong.

DR. DIGIOVINE: I had a question sort of, I guess it bridges this in usability. But what you're saying is that this is data that would be useful to health plans.

Do you have, do you know of any health plan that's gone out and looked at 100 of their patients who have a ratio that's less than .5 and could find a meaning in those, what those patients, was going on with those patients and could intervene and could actually improve compliance?

MR. HAMLIN: Not that I could name any names, no. But, no, again we have heard that they have been looking seriously at this measure because it has again been starting to be introduced in different programs for quality reporting as a replacement for what was existing.

And we're effectively waiting for those studies to be done and be published, if you will. And we do hear, and the lack of any other information we have not heard or hate mail for open lines that we get about measures when people are not happy with them.

So in the absence of any hate mail I'm assuming that people are still interested or intrigued and are still looking at this issue and

they're not going to just stay away because, you know, they don't think it's a valid measure.

DR. LAMPONE: Yes, in the information I've seen in reviewing this measure mainly what some of the plans looked at is when they had their members fall below that ratio threshold it would trigger an event of education to their membership and they would have programs with their providers to do this.

So it would help the provider target some of those members who are at risk for exacerbation based on that measure and try to get to a root cause of whether they were noncompliant or whether they were having other problems with meeting that measure.

They also made note that it also opens the conversation for the provider with that member to maybe explore other controlling medications that are easily administered, less multiple doses during the day and there was some data that looked at higher compliance rates when patients were able to be transitioned over to

more convenient controller medications.

So that's where I see the utility in some of those alternate measures that were not in this.

MR. HAMLIN: And I think we're perfectly in line with your vision.

CO-CHAIR BRATZLER: So other comments about validity? All right, we can go ahead and vote.

MS. AMIRAULT: Okay. Voting for validity for Measure 1799. Your options are one high, two moderate, three low and four insufficient. Again, validity for 1799.

(Voting.)

MS. AMIRAULT: Okay. Zero high, 15 moderate, five low and zero insufficient. And based on this we can move along.

CO-CHAIR BRATZLER: All right. We'll go to feasibility.

DR. COLLINS: Yes, as far as we know nothing has changed since the last approval. It seems highly feasible electronic data which is

already, you know, collected in mass as the Yoon study would show.

CO-CHAIR BRATZLER: Any questions or comments about feasibility? Okay, we'll go ahead and vote.

MS. AMIRAULT: Okay. Feasibility for 1799. Your options are one high, two moderate, three low and four insufficient. Again, feasibility for 1799.

(Voting.)

MS. AMIRAULT: Okay, 19 high, one moderate, zero low and zero insufficient. And based on the percentage we can move along.

CO-CHAIR BRATZLER: Usability and use.

DR. COLLINS: As far as usability it's publicly reported already in a number of places including consumer reports. I think the developers and we've had some nice discussions here already about the usability of this measure where the result means including some of the controversies around it. And I really don't have anything more to add.

CO-CHAIR BRATZLER: So I know you raised it earlier. But I do have concerns anytime we hold people accountable for a measure where there's a question in the evidence about the impact on outcome about unintended consequences of driving costs, use of medications just because we have a measure not because we're improving patient outcome.

So you've highlighted that before. But I just wanted to raise that one more time that does concern me some. Any other comments from the group, questions about usability? And this is a new measure so you haven't had time to see a lot of change over time. Okay. All right. I think no other comments we can vote.

MS. AMIRAULT: Okay. Usability and use for Measure 1799. One high, two moderate, three low and four insufficient. Again, usability and use for 1799.

(Voting.)

MS. AMIRAULT: Okay. Two high, 11 moderate, six low, and one insufficient. And

based on the percentage we can move along.

CO-CHAIR BRATZLER: And then the last vote will be on usefulness, what's the term just overall suitability, okay. So any other comments? Obviously we've highlighted, I think the developers have made it clear they are aware of the controversy around one study and are looking at that.

And the issue about questions of unintended consequences has been raised and some of the questions about validity. Anything else that we need to raise with the developers? Go ahead.

MR. HAMLIN: To address your comment about unintended consequences, these two measures are paired with the risk-adjusted relative resource use for asthma which do include specific subcategories for medication use and use of syncopation and outpatient services among 37 other things. So there is a way that we could track changes and resource use driven by these measures as a valued measure of this paired

quality.

CO-CHAIR BRATZLER: Almost a balancing measure. All right. Any other comments? So we'll take our final vote on this measure.

MS. AMIRAULT: Voting for overall suitability for Measure 1799. One for yes and two for no.

(Voting.)

MS. AMIRAULT: Okay, 12 yes and eight no, grey zone.

CO-CHAIR BRATZLER: All right, thank you. So we'll let our friends from NCQA give us one more introduction to the next measure which is 0577. This is Spirometry Testing and Assessment and Diagnosis of COPD. And Dr. Lang is back at the table now to be a participant in the conversation.

MS. ROTH: So again, this is a health plan level measure. It assesses whether patients 40 and older with a new or newly active diagnosis of COPD had spirometry testing to confirm their diagnosis.

And plans calculate this measure by identifying people with a COPD diagnosis during the measurement year and then excluding those people who had a COPD diagnosis any time in the two years prior to that. So then they arrive at a population who has new COPD.

And we did test different options for this time interval including looking one year back, looking only six months back for the COPD diagnosis history. But we found that the two years was the most specific at getting to the newly diagnosed patients.

The measure is based on guidelines and evidence that spirometry should be performed to diagnose airflow obstruction in patients with respiratory symptoms in order to make a clinical diagnosis of COPD and to determine appropriate therapy. We did cite several studies from AHRQ evidence review that found that spirometry to confirm diagnosis helps providers identify those patients who might benefit from medication that would lessen their risk of COPD exacerbations.

But there is some more research needed to assess the link to outcomes. And then we did also provide some data following the Work Group call showing that the measure rates have improved six percentage points since 2008. And this was across all the plan types, commercial, Medicaid and Medicare.

But although there has been some improvement rates do continue to be low. So there's still opportunity for improvement going forward.

CO-CHAIR BRATZLER: Thank you. So our discussants are James and Todd.

DR. DORMAN: I think that leaves me.

CO-CHAIR BRATZLER: Okay.

DR. HARRIS: I'm on the line.

DR. DORMAN: Okay, wonderful. Do you want me to start or --

DR. HARRIS: Go ahead.

DR. DORMAN: Okay. It might be easier that way. So this is a, I think thank you for the explanation and I think that sets the stage

well. This is really about the use of spirometry testing in the diagnosis of COPD.

It's a process measure at the level of analysis of the health plan first endorsed in '09 and then endorsed in 2012 and now up as a maintenance measure again. Under the level of evidence, new evidence updated evidence was provided.

This included from the GOLD Guidelines which references over 600 studies, a recommendation additional work through ACP\ATS and European Respiratory Society which also made the same basic recommendation that spirometry should be obtained to diagnose air flow obstruction in patients with symptoms and cites additional references of controlled trials meta-analysis and systematic reviews.

And then there's a third guideline that is also listed. And these are all updated so there is new evidence. I think the discussion we had centered around the six month to two years and did not see hard evidence of this but felt

that there was face validity and appreciated the explanation and felt that the vector of the evidence was not impacted by the new data.

And in fact we thought it was enhanced and merely the magnitude of that vector was even stronger so that the evidence base grew for this recommendation.

CO-CHAIR BRATZLER: So this is a maintenance measure. James, did you have any other comments about that?

DR. HARRIS: No, I think, you know, again one of the things that the, in the initial endorsement there was some question about sort of the link between the process and the outcome. But some of the additional studies are providing a little bit more of that link.

So again, I think it's a little bit stronger than it was previously.

CO-CHAIR BRATZLER: So based on that conversation, there is new evidence but it just seems to make the, strengthen the argument from the previously endorsed measure. So raise your

hands if you think we need to revote the evidence at this time.

Okay. Seeing none we'll go on then with our discussion of --

DR. DORMAN: Thank you, James. You want me to continue or --

DR. HARRIS: I'll take this one if you want. So this is the gap, performance gap study. Again this was one of the things I think that we had some questions about in terms of looking at some of the numbers that were provided initially from HEDIS in terms of looking at the breakouts from commercial plans and Medicare and Medicaid.

And didn't see, there certainly are sort of differences between the lowest being the Medicaid plan and the highest being commercial, you know, with 42 percent Medicaid being 31. So it certainly does look there's some difference in terms of entities, opportunities to improve there.

There was no information on demographics, race and ethnicity provided. But

again, you know, because of the difficulty in breaking that information out in terms of the data that's actually collected in this measure.

CO-CHAIR BRATZLER: Todd, anything else?

DR. DORMAN: Nothing to add.

CO-CHAIR BRATZLER: Any other questions or comments about gaps? Okay, we can go ahead and vote.

MS. AMIRAULT: Okay. Performance gap for Measure 0577. Options are one high, two moderate, three low and four insufficient. Again, performance gap for 0577.

(Voting.)

MS. BAL: Could everyone vote one more time in the room? We're missing a few more people. Thank you.

MS. AMIRAULT: Okay, nine high, 12 moderate, zero low and zero insufficient.

DR. DORMAN: So, James, I'll do reliability and let you do validity. Is that okay?

DR. HARRIS: Sure.

DR. DORMAN: Give me one second. So electronic clinical data, claims data specifications were not updated since the last review. This is a measures score. Beta-binomial method was used for signal-to-noise and they were quite high across multiple populations, .92 to .97.

And give me one second. Let me get to the other part of my notes. So I think the discussion centered mostly around again the six month to two year period of time. And there was some discussion around and maybe it's worth reiterating a piece here or having the developers respond on the contrasting evidence between spirometry reading to improve care and confirming diagnosis.

And I just want to add I think this comes about in part related not only to trying to drive towards quality but by the title of this particular one which says in assessment and diagnosis when the evidence is all about the

diagnosis. And the word assessment I think leads some people to believe ongoing management.

James, do you want to add anything on reliability?

DR. HARRIS: No, nothing outside that.

CO-CHAIR BRATZLER: So NQF staff tell me the specifications really have not changed, that we don't have to vote on this unless you want to. So how many think we need to vote on reliability?

Okay. We'll move on to validity. And I guess Mitch is, you prefer Mitch?

DR. HARRIS: Yes, James, Mitch is fine. I answer to both. My parents played a cruel trick on me when they named me one thing and called me the other. So substitute teachers were always fun at school.

So in terms of validity there is some change here in terms of when the measure was initially examined and endorsed. Initially it was just face validity was provided. Since then sort of there have been some updated validity

testing using some other and comparing to other measures.

So this is a place where before the validity could have only been rated medium and now it's eligible for high. But I think again you do see, I think some strong metrics here in terms of the comparisons to some of the other measures and the appropriate direction or the direction that they hypothesized them being.

There are no exclusions for this measure and there is no risk adjustment. So that maybe something to think about and then in terms of missing data as well they just cite sort of the audit processes that are in place currently to look at those, to look at the measure that would sort of help alleviate any of the concerns that might be around missing data.

CO-CHAIR BRATZLER: Todd, anything else?

DR. DORMAN: Nothing to add.

CO-CHAIR BRATZLER: All right. And in conversations about validity, questions about

validity? So we do need to vote on validity. So we'll go ahead.

MS. AMIRAULT: So validity for Measure 0577. Options are one high, two moderate, three low and four insufficient. Again, validity for 0577.

(Voting.)

MS. BAL: Could everybody just vote, never mind. We got it. Thank you.

MS. AMIRAULT: Eight high, 13 moderate, zero low and zero insufficient. And based on the percent we can move on.

DR. DORMAN: Mitch, in the spirit of alternating I'll do feasibility and you've got the last part usability, okay?

DR. HARRIS: Sure.

DR. DORMAN: So feasibility all the fields are defined in electronic claims with differing individuals that obtain the data and code the data the NCCU conducts independent audits. And I don't know that there's much else to say.

CO-CHAIR BRATZLER: Any other questions, concerns about feasibility? Hearing none we'll go ahead and vote.

MS. AMIRAULT: Feasibility for Measure 0577. One high, two moderate, three low and four insufficient.

(Voting.)

DR. NISHIMI: Can, point there, got it, okay.

MS. AMIRAULT: Sixteen high, five moderate, zero low and zero insufficient. With this percentage we can move along.

DR. DORMAN: Okay, Mitch, go talk about usability.

DR. HARRIS: Sure, thanks. So I think usability again a couple of things to look at. It is currently being used and the NCQA's Data Health Care Annual Report and also through reporting consumer reports used in Quality Compass.

So there are a couple of places where it's already being examined and used for

benchmarking and quality improvement. I do think one of the things, again some additional information was provided but initial data sort of, you know, there was some concern about is there actually improvement being shown.

The rates broken out by the plans were pretty stationary. But some of the additional data provided went back I think into 2008 you could see that there's, you know, in some cases five to ten percent sort of increases between the years the data that they provide.

So there is some change going on over time. But as we mentioned in the gaps also there is a lot of room for improvement as well.

DR. DORMAN: I wouldn't, nothing to add.

CO-CHAIR BRATZLER: All right. Thank you. So any other discussion about usability? All right. Seeing none, let's vote.

MS. AMIRAULT: So for usability and use for Measure 0577, one high, two moderate, three low and four insufficient.

(Voting.)

MS. AMIRAULT: Okay, seven high, 13 moderate, one low and zero insufficient. And based on the percentage we can move on.

CO-CHAIR BRATZLER: All right. And our last vote will be on overall suitability. Any other conversation, comments for the developer that we want to send back?

Everybody is quiet. They're waiting for the break. So we'll go ahead and vote on suitability.

MS. AMIRAULT: Okay. Overall suitability for Measure 0577, one for yes and two for no.

(Voting.)

MS. AMIRAULT: Twenty-one yes and zero no for 100 percent.

CO-CHAIR BRATZLER: All right. Thank you very much. So I think the last thing we have to do before the break is we've just covered three metrics that look at pharmacotherapy for asthma.

Two by NCQA and one by the American Academy of Allergy and, Asthma, Allergy and Immunology. So --

DR. NISHIMI: So in this case the Committee can encourage harmonization. They may decide that they prefer one over the other. I would note that or one is over the other.

I would note that they are at different levels of analyses. So that's often a justification provided by the Committee and they do use different data sources. The AAAAI is a clinician level measure and obviously the two NCQA measures are plan level.

CO-CHAIR BRATZLER: And the one comment I would make and we talked about this extensively I think our major developer heard that was that for the AAAAI measure the performance rate combined control medications is very, very high on the metric.

And we made the recommendation to consider reporting this in different ways. And the ratio measure actually that NCQA has

potentially could be, it would be more difficult but there may be some way to report a ratio measure as a part of that or have as we discussed multiple numerators for the AAAAI measure also.

DR. NISHIMI: So the instruction would be to the developers in that regard is to have that kind of discussion. Is there anything else that the Committee wants to have them discuss?

DR. O'BRIEN: So I appreciate the fact that the ones at a clinician level and the other at the health plan levels. It would be great from an alignment standpoint of actually improving patient care.

There was alignment in those two measures so that you could go to the provider and say this is what you need to do for the health care plan to be successful as opposed to the possibility of them being disparate measures.

DR. OHTAKE: I just have a question about the count you made about 0047. It sounds like and well what we read is in the numerator it's just all controller medication, not

separating out corticosteroids.

But yet in the very first preamble in the brief description it says that three rates are reported where they do inhaled corticosteroids non-ICS and then everyone. I was wondering if the --

MS. SWAIN-ENG: Yes, there are three separate reads that they do report out. I think his comment though was to actually report out on because the list was so long perhaps you would want to make a smaller list. I think that was the comment if I'm correct.

CO-CHAIR BRATZLER: Well and the other thing though is the metric that actually gets used for public reporting, the value modifier and everything else is the combined rate which is very, very high rates of performance.

And I'm not convinced that is the most appropriate metric that we want to drive. So that's what gets used for all of these payment and reporting programs for 0047. So and that's where I think the ratio is an interesting concept

of looking what you're considering controller medications to all other asthma medications.

DR. NISHIMI: Anything else? Okay. So the staff writes that up and then the developers will discuss and when it comes back for maintenance we'll see what happens.

CO-CHAIR BRATZLER: So you guys were incredibly efficient. It looks like we either gained five minutes of break or we can come back at ten after. What do you prefer? What's that? All right, 3:10 we're starting again.

(Whereupon, the above-entitled matter went off the record at 2:55 p.m. and resumed at 3:11 p.m.)

CO-CHAIR BRATZLER: All right, very good. We did good, so let's keep moving forward. We have three more measures to discuss this afternoon.

The first one is 0091, COPD Spirometry Evaluation from the American Thoracic Society. With us here in person is Sue Frechette and Bela Patel is on the phone, and they'll give us the

brief introduction to the measure.

MS. FRECHETTE: Thank you. I'm going to try to make myself heard. Bela, are you on the call at this point?

(No response.)

MS. BAL: Operator, could you make sure that Bela Patel has access, has a speaking line? Thank you.

MS. FRECHETTE: So Dr. Bela Patel who is our COPD expert who will be joining us shortly, if she's not on the call yet. So let me just give you a brief introduction.

First of all, the ATS spirometry measure was originally developed by the ATS in conjunction with the AMA-PCPI and has been in use in the PQRS program since 2007.

It was initially endorsed by the NQF 2009, and was recently reendorsed in 2012. It's also been in use for the ABIM MOC program for the past three years. In late 2014, stewardship was transferred from the AMA-PCPI to the ATS. So we have recently taken over stewardship of this

measure.

Performance since 2008 shows a favorable trend, and that trend is well-aligned with the literature. Just as a side note, in your work sheet the numbers from 2008 were reported. I think they're approximately 50 percent performance. There was some additional information. It's on page 16 of the full application.

But just to quickly summarize, the trend since 2010, 2010 performance was up to 56 percent. The most recent numbers were from 2014, and performance was at 67 percent. So performance is definitely moving in the right direction on this measure.

We believe strongly that accurate diagnosis of COPD is essential to provide high quality COPD management, and as you know that will improve quality of life, reduce exacerbations and ultimately reduce the cost of care. I won't repeat the details. We have a large amount of information in our application so

that information's all there.

But I'd just like to point out that it's critical that patients who are undiagnosed with COPD may not benefit from medication that could improve their lives, reduce costs, etcetera.

But there are also many patients who do not have COPD, that when they receive COPD medication have negative consequences, and that also may not benefit from medication that could ultimately improve whatever condition, respiratory condition they may have if it's not COPD.

So that's why we think spirometry is absolutely critical for diagnosis, for treatment. So based on trend, the literature, our belief in the importance of accurate diagnosis, we believe that this measure continues to be a good measure, and we would seek reendorsement if you're in agreement.

There was a brief discussion at the workshop meeting about the clarity of the

numerator for this measure. We believe we can add a little bit of clarity and are glad to discuss it further in the numerator details, which is Question No. S-6 in the full application that ought to help with this discussion.

So with that said, Dr. Patel and I welcome the discussion and are available to answer any questions you may have on this measure. So thank you.

CO-CHAIR BRATZLER: And it's my understanding Dr. Patel is on the line. Can you hear us now or can we hear you?

DR. PATEL: Yes, I can hear you now. Hopefully you can hear me.

CO-CHAIR BRATZLER: Yes, we can. Did you have anything to add to that brief introduction?

DR. PATEL: No, I do not.

CO-CHAIR BRATZLER: Very good. All right. So our discussants are Gerene and Edgar.

DR. BAULDOFF: Okay. The COPD spirometry evaluation measure is a -- the

description is the percent of patients at 18 or greater years of age with a diagnosis of COPD or who have had spirometry documented. It's a process measure that is in consideration for maintenance endorsement, and it's at -- the level of analysis is at the clinician group or clinician team.

The developer has provided systematic review of the evidence, including quality, quantity and consistency, as well as grading of the evidence. There's been no change in the evidence since the prior approval.

The evidence algorithm indicates a high rating, and that the work group had reported that we thought there was no need to discuss the evidence, even though there has been an improvement in the gap. But the gap still is at greater than 30 percent.

CO-CHAIR BRATZLER: Any --

DR. JIMENEZ: No, nothing.

CO-CHAIR BRATZLER: All right. So we have a previously endorsed measure, strengthening

of the evidence but otherwise no real change since the original endorsement in '09 and '12. James, did you have a comment?

DR. O'BRIEN: Yeah. Just a quick question for the developers. I was unclear if this is patients who -- with a diagnosis of COPD who have ever had spirometry or that they repeatedly get spirometry documented?

MS. FRECHETTE: It's ever had spirometry. Ever.

DR. O'BRIEN: Ever.

CO-CHAIR BRATZLER: Any other questions or comments? So I'm going to ask anyone to raise your hand if you think we need to vote on the evidence.

(Pause.)

CO-CHAIR BRATZLER: Seeing none, we'll go on with the conversation about performance gaps.

DR. BAULDOFF: Do we need to vote on gap? Is that what we're discussing. Okay. The gap was reported at 67 percent. It is improved

from about 50 percent over the last four or five years.

Those were from PQRS scores, and that it's reported that the gap remains limited by current literature. There's just really nothing new. There's a paucity in the literature.

DR. JIMENEZ: There will be a slowdown in the gap also because MOC has been put on hold. I don't know if you are aware of that too, right, for at least a couple of years now that we understand.

DR. BAULDOFF: There were a couple of things that came out of the work group, that there was a couple of articles that were mentioned out of the work group. One was that dealt with individual FEV1 trajectories, the differential trajectories and other was clinical inertia in the management of chronic COPD. These would just supportive of the measure.

CO-CHAIR BRATZLER: So I'm curious. The data that the gap is reported on, is that strictly from the ABIM or is that national?

MS. FRECHETTE: I'm sorry. The data is from CMS. It's from 2010 through 2014.

CO-CHAIR BRATZLER: Okay. So that's reported at the level of the practitioner. So there really isn't any ethnicity, racial, you know, other stratification data out of that that you can --

MS. FRECHETTE: That's correct. We've requested it, but they don't striate the data at that level.

CO-CHAIR BRATZLER: Okay. Any other questions then or comments about gaps?

(No response.)

CO-CHAIR BRATZLER: If none, we'll go ahead and vote.

MS. AMIRAULT: Voting for performance gap for Measure 0091, 1 for high, 2 for moderate, 3 for low and 4 for insufficient.

(Voting.)

MS. AMIRAULT: Okay. 10 for high, 9 moderate, 1 low and 0 insufficient, and based on the percentage, we can move along.

CO-CHAIR BRATZLER: All right. We'll discuss reliability.

DR. BAULDOFF: Okay. From a reliability standpoint, the specifications have not been updated since the previous review, and the reliability testing algorithm rated as high. There is some confusion, which you mentioned Sue, on the work group all around the time window on spirometry.

Is it one year, within the past one year or any time, and how do we manage multiple measurements were the questions that had come up.

MS. FRECHETTE: I forgot to turn this on. The clarification is during, and this is -- it's probably written better than I can express it on the response to S-6 in the specifications. But the intent is that once a year look back to confirm that the patient has had at least one spirometry test confirming that the patient has COPD.

So it is possible that there could be an additional test that's done that perhaps is

not necessary. But based on the literature and based on the research that we've done on these measures or this measure, under-use of spirometry is a much greater issue than over-use at this point.

So we realize it could be a couple of over-use data elements sneaking into the overall performance, but we don't think that's a big issue at this point.

DR. JIMENEZ: The time limit was done on initial part of the measure, right? I mean it was any time. It was spirometry at any point in time.

MS. FRECHETTE: Spirometry at any point in time. The intent is that COPD is confirmed before a patient starts getting COPD-specific medication.

CO-CHAIR BRATZLER: Any other questions? Yes, Bruno.

DR. DiGIOVINE: Just in terms of data capture, what are the instructions around if this is being abstracted on -- is the intent that the

physician should always at least once a year mention PFTs that were done in the past? Is that the sort of the baseline request?

MS. FRECHETTE: The intention is that once a year, the physician just notes that yes, the COPD has been confirmed before I treat or change treatment or make adjustments to therapy.

DR. DiGIOVINE: Can you clarify that, because maybe I read it wrong? It said that the note has to have the FEV1 and the FEV1 VIFs, the FVC ratio in it unless I read it wrong. So that would not be merely writing a note that confirms that I previously had tested the patient and they were confirmed to have the diagnosis of COPD. That would be having to recite the results of that previous test. Am I misinterpreting?

MS. FRECHETTE: I can go back to the details, but I believe it's to confirm that somewhere in the chart it's noted that the spirometry testing was done and what the results were, which includes the ratio and the FVC1. Does that clarify?

CO-CHAIR LANG: So let me follow up on that. So then if Bruno sees a patient at Henry Ford Hospital in Detroit with mild COPD, and that patient relocates to Cleveland and sees me at the Cleveland Clinic, and I mention that the patient has had a spirometry in the past done in Detroit, what is the attribution, if this is at the clinician level?

MS. FRECHETTE: Well, that's one of those things where it depends. It's a complex situation. Very often a physician or a patient will be diagnosed with COPD by one physician, treated by another physician. The intent is not to repeat the spirometry; it's to confirm the spirometry was done and what the results were. There's always going to be patients that are going to move from one doc to another.

CO-CHAIR LANG: So I mean the age is another issue that I'll raise, but I mean your age is 18. So if that patient has had a spirometry, then that patient, wherever he or she goes for the rest of his or her life has

fulfilled this measure? I mean is that -- am I interpreting this correctly?

MS. FRECHETTE: That's the intent at this point.

DR. DiGIOVINE: I guess just to follow up on Todd's point, I guess -- I think it sounds like we would agree it would be more in line with what you're trying to get at, if the clinician states the patient had a spirometry which was consistent with the diagnosis of COPD, without actually adding the actual FEV1 or the ratio, that that would meet your GOLDs as a quality measure?

MS. FRECHETTE: Yes.

CO-CHAIR BRATZLER: Richard.

DR. MURRAY: Just to follow on with-- just to follow on, if once at age 18 is enough for the rest of your life, then the intention must be for accuracy in the diagnosis and in no way for a medical assessment over time; is that right?

MS. FRECHETTE: I'm sorry. I didn't

hear the second half of your question.

DR. MURRAY: The intent of the measure is to drive accuracy in diagnosis at the origin, not for assessment over time?

MS. FRECHETTE: Correct. It's to confirm COPD as opposed to confirm or to check whether or not it's changing.

DR. MURRAY: Okay.

TL So I just had a clarification. On the cohort of patients you're following, they're enrolled in a registry; is that correct?

MS. FRECHETTE: This measure is set up to be reported either for a registry or through claims.

TL Or through claims. So when patients get this spirometry done 18, 19, 20, and they move and they go out of state or maybe even on an exchange plan, how do you -- how do you follow that trail and has that been problematic?

MS. FRECHETTE: I don't know if Bela has a more definitive response. My response would be that when a patient moves somewhere far

away or to a completely different treatment area, some tests will be repeated. That's just a reality of how we practice medicine in the U.S.

But the intent is if they've had a test, they wouldn't have to repeat the test or regular repeat the test.

DR. LAMPONE: So how does a provider wind up participating in the registry if you have a patient that started out in Texas and winds up in Iowa?

MS. FRECHETTE: Well, the intent -- again the intent is that a physician would not treat somebody for COPD without confirming that they do have COPD with spirometry. So that does not require retesting a patient. What it requires is confirming that the test was done.

So hopefully in this age of electronic medical records, we can share those documents. Again, that's a little bit futuristic. They should be sharing documentation and sharing test results.

DR. PATEL: This is Bela. Now if you

were a physician treating this patient, either you would be able to get ahold of those prior spirometry results and review and confirm in your medical record, or like a lot of -- like infrequently, it doesn't happen infrequently in which the practitioner's unable to get the results and they would repeat the spirometry.

CO-CHAIR BRATZLER: So this is a PQRS measure that can currently be used. I can tell you that we've run the report before. We have a structured field documentation of the FEV1 and the FEV to F for spiral capacity ratio.

Once that's in there, I can run a report for any doctor at any point and see whether there's structured field documentation of those numbers or not, and it lives in perpetuity in those structured fields.

MS. WEST: I guess I'm just trying to better understand how you basically keep that patient from being counted multiple times over multiple years, if they're already getting credit 18 years ago when this patient was initially

diagnosed.

Are you now counting that patient every single year that you're submitting the PQRS data and getting credit for doing something that you did years ago? Like I guess I'm just trying to wrap my head around how you're doing accountability.

MS. FRECHETTE: Well, I think my perspective is what you call credit is doing the right thing, and so yes, year after year you can report that yes, I can confirm that I'm treating this patient with COPD medication, and I've confirmed that they have COPD.

DR. DiGIOVINE: I just want to -- just wanted to sort of echo in support. This is why we -- this is sort of a point of how bad we are, is that despite all of these sort of easy ways to be compliant, our compliance rates are still 40 percent or 50 percent. I mean we -- we do a lousy job at this, which is I think the importance of having some measure of this.

MS. FRECHETTE: And that's 100 percent

aligned with the research literature as well.

CO-CHAIR BRATZLER: So I'm going to -- unless anyone else has a burning comment, recommend we go ahead and vote on reliability.

MS. AMIRAULT: Okay. Reliability for Measure 0091, 1 for high, 2 for moderate, 3 for low and 4 for insufficient.

(Voting.)

MS. AMIRAULT: Okay. 4 high, 15 moderate, 3 low and 0 insufficient, and based on the percentage we can move forward.

CO-CHAIR BRATZLER: Validity.

DR. BAULDOFF: Okay, validity. There was a new face validity that was conducted by the ATS Clinical Practice Group. It was a 91.7 percent agreement, at agreed or strongly agreed, that the measure can distinguish good from poor quality. Both measure score and data element testing against the gold standard were reported.

The meaningful difference is reported to be 45.6 percent to 47.1 percent by the group size, the provider group size. There's -- it was

still unclear the work group regarded reported usage in gap is the spirometry to confirm diagnosis or does this focus on routine spirometry use?

Clearly, it focuses on routine spirometry use. This was not specific to just finding people being diagnosed, correct? It could be, but it wasn't limited to that.

MS. FRECHETTE: Correct. The intent is not routine spirometry use; it's a onetime confirmation.

DR. BAULDOFF: Okay, thank you.

CO-CHAIR BRATZLER: Any other questions or comments about validity?

DR. BAULDOFF: The validity algorithm came up as moderate.

CO-CHAIR BRATZLER: Other questions or comments?

(No response.)

CO-CHAIR BRATZLER: All right. We'll go ahead and vote on validity.

MS. AMIRAULT: Okay. Validity for

0091. The options are 2 moderate, 3 low and 4 insufficient.

(Voting.)

DR. NISHIMI: Okay. Everyone point again.

(Voting.)

MS. AMIRAULT: Okay. 0 high, 18 moderate, 3 low and 0 insufficient, and based on the percentage we can move forward.

CO-CHAIR BRATZLER: Okay. Feasibility.

DR. BAULDOFF: From a feasibility standpoint, the work group and the reviewers, this appears a reasonable measure, as the information is collected and generated during routine care and it's available electronically.

DR. JIMENEZ: This is just to emphasize in what Bruno was saying a while ago, is as a comment, you know, that how -- what we generally do, if we have 1 in 3 physicians, I mean, they'll never use this and being so simple to apply. It's really feasible.

So it is something that, you know, it is alarming. If we have this the aviation industry, we would have airplanes dropping in the street every day.

CO-CHAIR BRATZLER: We're not high reliability you're saying? Any other comments about feasibility?

(No response.)

CO-CHAIR BRATZLER: If not, we'll go ahead and vote.

MS. AMIRAULT: Okay. For Feasibility for 0091, 1 high, 2 moderate, 3 low and 4 insufficient.

(Voting.)

MS. AMIRAULT: Okay. 12 for high, 9 moderate, 0 low and 0 insufficient, and based on the percentage we can move along.

CO-CHAIR BRATZLER: Use and Usability.

DR. BAULDOFF: Usability, this data is publicly reported. According to the application, the plan is to include in the CMS Physician Compare Program. It wasn't included as of

December 14th, 2015. It is in the PQRS. Has it shown up in the CMS Physician Compare yet? I don't think that's changed again yet, has it?

MS. FRECHETTE: I have not checked yet.

CO-CHAIR BRATZLER: Okay. Physician Compare currently I think only has the metrics that are reported via the web interface for PQRS. Therefore, this is not one of the web interface measures. But moving forward, it could show up you would think. Any comments about use or usability?

(No response.)

CO-CHAIR BRATZLER: If no, we'll go ahead and vote.

MS. AMIRAULT: Okay. Usability and use for Measure 0091, 1 high, 2 moderate, 3 low and 4 insufficient.

(Voting.)

MS. AMIRAULT: If everyone could just point one more time? Sorry.

(Voting.)

MS. AMIRAULT: Okay. 8 high, 12 moderate, 1 low and 0 insufficient, and based on the percentage we can move along.

CO-CHAIR BRATZLER: All right, and our last discussion will be on overall suitability. So I think the primary issue I heard discussed frequently is that the bar is pretty low on this metric. Has spirometry ever been done and documented in the electronic -- in the record.

I think there was a general consensus that that's probably too low of a bar, but there's still a big gap in performance. Other than that, are there other issues that we need to raise with the developer?

DR. BAULDOFF: The only other one is probably outside of this, and that's around related and competing measures, and we can discuss that later.

CO-CHAIR BRATZLER: Okay, very good. Yes.

MS. WEST: I guess part of this is related to kind of the bar being set low. I

thought that there was kind of a general consensus that checkbox measures were kind of being thrown out to them.

CO-CHAIR BRATZLER: I'm not sure this is completely checkbox though. This actually requires you to do a formal test and then document it in the chart. But you have to formally do the test so --

MS. FRECHETTE: If I can add, you don't have to be the one to do the test. You need to know what the results are, and at this point, 67 percent of patients have it. That means, you know, 30 percent don't have the test done. That means those people who have unconfirmed diagnoses may be getting erroneous medication and probably not benefitting from the medication.

CO-CHAIR BRATZLER: And that 67 percent, again I'll highlight, is those people who voluntarily decided and chose that measure to report because PQRS. It's up to you to pick what you want to report oftentimes. So if it was

broadly required, it could be quite low.

MS. FRECHETTE: The literature shows that spirometry testing is done on between 50 and 60 percent of patients. So this is your point. The research literature shows it's even lower.

CO-CHAIR BRATZLER: All right. Any other comments about overall suitability?

(No response.)

CO-CHAIR BRATZLER: Okay. Let's go ahead and vote.

MS. AMIRAULT: Okay. So voting for overall suitability for Measure 0091, 1 for yes and 2 for no.

(Voting.)

MS. AMIRAULT: Okay. 19 yes, 2 no. Based on 90 percent, we can move on.

CO-CHAIR BRATZLER: Okay, very good. Thank you very much. So the next measure keep Sue at the table and Dr. Patel on the phone is No. 102, COPD inhaled bronchodilator therapy from the American Thoracic Society. If you can give us a brief overview.

MS. FRECHETTE: Okay, the background on this measure, this bronchodilator measure was also developed by the ATS working with the AMA-PCPI and in use since 2007. Initially endorsed by NQF 2009, most recently endorsed 2011. It was added to the MRC COPD program in the fourth quarter of 2015. So it's a part of that program, and also in 2014 stewardship was transferred from the AMA-PCPI to the ATS. So we were just taking over stewardship of this measure.

Performance since 2008 shows a favorably trend in high performance above 95 percent. However, this is not aligned with the research literature that shows appropriate medication management in approximately 60 percent of our patients, COPD patients.

So the ATS Subcommittee on Measures did quite a bit of work, quite a bit of research to try to determine why this misalignment existed. What the committee found is that inhaled bronchodilators do include both short and long-acting bronchodilators. The evidence-based

guidelines, especially the most recent guidelines, clearly state that long-acting bronchodilators are preferred over short-acting bronchodilators. They have better outcomes and particularly around reducing exacerbations.

The subcommittee also noted a transcription error from the updated specifications in the 2013 annual review. So for these reasons, the trending, this misalignment with the research literature and to correct the error, the subcommittee recommends a change in the numerator specifications.

So this is just to point out that our entire application is on -- includes the new specifications, which specify long-acting inhaled bronchodilators. So the only change to the current specifications are the addition of long-acting, not all inhaled bronchodilator.

This change was approved by CMS for use in the 2017 PQRS performance year. The committee believes that as we make this -- as we implement this change, performance will go down

significantly if we're looking only at long-acting bronchodilators.

So there was a discussion at the workshop as to whether or not this measure was topped out. We believe this measure is not topped out. We believe that many patients are receiving short-acting bronchodilators when they would benefit from longer acting bronchodilators.

Note that in our application as we sent, the application is all on the new language. The exceptions obviously are that the performance data for the past six years, as well as on the tests, where tests were done on the current measures.

So again, Bela Patel and I welcome the discussion and are ready to answer any questions you may have.

CO-CHAIR BRATZLER: All right. Dr. Patel, did you have anything to add to that?

DR. PATEL: No. Sue did a great summary.

CO-CHAIR BRATZLER: Okay, very good.

So our two discussants are David and Gerene.

CO-CHAIR LANG: So this is a process measure at the level analysis of group practice, clinician. The numerator, as Sue stated, the numerator stipulates patients who are prescribed the long-acting bronchodilator, the denominator being all patients aged 18 and older with a diagnosis of COPD based on symptoms and/or spirometric criteria.

The evidence has been updated based on three clinical practice guidelines, with diagnosis and management of chronic disruptive lung disease and I don't have any other concerns regarding the evidence, which I don't believe we need to address any further.

DR. BAULDOFF: The evidence algorithm was high. That was the only other thing.

CO-CHAIR LANG: I'm sorry?

DR. BAULDOFF: The evidence algorithm was high. That was the only other point.

CO-CHAIR BRATZLER: Which is comfortable, we did change the numerator.

CO-CHAIR LANG: Well, the numerator has been changed. The numerator stipulates patients who were prescribed the long-acting bronchodilator.

DR. DiGIOVINE: I recognize that, since the previous reports. It's a good question.

DR. NISHIMI: Not if you feel the underlying evidence.

CO-CHAIR BRATZLER: Okay. So raise your hand if you think we need to re-vote on the evidence for this measure.

DR. O'BRIEN: I just have question to clarify. I was just confused reading through it. Are inhaled corticosteroids considered compliant with this measure?

MS. FRECHETTE: That's a question for Bela.

DR. PATEL: No, they are not. Just the inhaled bronchodilators. Not inhaled corticosteroids.

DR. O'BRIEN: So actually I was going to stipulate that -- so what you're referring to

here are long-acting beta agonists or long-acting muscarinic agents?

DR. PATEL: Correct.

CO-CHAIR LANG: So a major issue here for us to discuss, which Sue -- I'm sorry, moving on to performance gap.

CO-CHAIR BRATZLER: So let's make a decision. Are we going to vote on evidence for -- since we've with those -- Edgar, do you have a comment?

DR. JIMENEZ: I just noticed here in the notes from the reviewers, there was one -- does the committee agree there's no need to repeat discussion on both, on the evidence and it says yes. I mean so someone -- in the notes that we have.

So my question is, I mean with the new evidence, I would say it would probably be appropriate to vote on the newer evidence, right, that's long-acting, that I mean really what we're looking at, because there has been a little change in what we had from the previous approval.

CO-CHAIR BRATZLER: James.

DR. O'BRIEN: And this may be, I guess, for the staff. At what point does a maintenance measure whose numerator has changed actually become a new measure? What are those criteria by which we actually use different -- because one could imagine taking something that's existing under an NQF number and wholly changing its specification, but acting like it's maintenance.

DR. NISHIMI: So whether there's a material change or not is actually something that NQF is grappling with right now and putting limits, if you will, around. So for purposes of this review, you can decide that, you know, this -- that you need to discuss the evidence.

Obviously, you would discuss it under specifications. But we did not assign it a wholly new number. It was -- because of the flux, frankly, in what is material and what isn't.

CO-CHAIR BRATZLER: So I don't want

the conversation about whether we should vote. It would take longer than just voting so should we just go ahead and vote on the evidence? Let's vote on the evidence and then move on.

MS. AMIRAULT: Okay. Voting for evidence on measure 0102, 1 for high, 2 for moderate, 3 low and 4 insufficient.

(Voting.)

MS. AMIRAULT: Okay. 9 high, 12 moderate, 1 low and 0 insufficient. Based on the percentage we can move on.

CO-CHAIR BRATZLER: Okay. Performance gap.

CO-CHAIR LANG: Performance gap. So for those of you who are sufficiently fortunate to have the original document either in your hands or on your screen, I would refer you to page three. The developer stipulates that in the 2008 data, 54 percent of patients reported on did not meet the measure.

And then furnished data for 2012 to 2014, which described the narrowing of this gap

to 1.5 to 4.1 percent, and that's what Sue described in terms of the obfuscation, at least by the fact that individuals who received prescriptions for short-acting bronchodilators were included in the numbers I just mentioned.

So I believe what the developer is advocating is moving forward with the understanding that these numbers don't apply to the newly-stipulated numerator of inhaled, long-acting bronchodilators, either long-acting beta agonists or long-acting muscarinics.

However, there are no data we have regarding a performance gap pertaining to that numerator. Is that correct?

MS. FRECHETTE: The numerator being for long-acting?

CO-CHAIR LANG: Yes.

MS. FRECHETTE: Not yet.

CO-CHAIR LANG: So we have no performance -- no data regarding performance gap.

DR. PATEL: Correct.

DR. DiGIOVINE: I'd like to, if I

could make a comment. I think we -- no one else would bring it up but now. So we are now saying that anybody with an FEV1 less than 60 percent of predicted requires a long-acting bronchodilator. The GOLD classification says patients who are gold stage A, which includes patients who have FEV1s down as low as 50 percent. The preferred therapy is a short-acting beta agonist or muscarinic agent.

So I need some understanding on why it would be a quality measure to say everybody with an FEV1 less than 60 percent or predicted, even though whose are not symptomatic, would require a long-acting beta agonist or a muscarinic agent?

DR. PATEL: So GOLD weight classification requires a FEV1 of greater than 80 percent; correct? I don't have the document in front of me.

DR. DiGIOVINE: Yeah. So gold A, I don't know if you wrote gold, Dr. Patel, but it's not -- it is confusing for all of us. So GOLD A includes patients who are spirometric class

either GOLD 1 or GOLD 2. GOLD 2 goes down to 50 percent of predicted.

DR. PATEL: Correct, and so the committee felt that since the guidelines had not only COPD measures but also included symptomatology, that based on the published, randomized trial that it was now appropriate -- there's enough evidence to support long-acting bronchodilators.

Because when we looked at -- if we didn't have to include the symptoms, then perhaps we may have overshot that guideline. But because you have to have symptoms and the specifications, we thought it was appropriate.

CO-CHAIR BRATZLER: Okay. I'm not sure I quite understood that.

DR. DiGIOVINE: Sorry. I'm just -- my point was just that I think the way the COPD guidelines are written, the patients who have no symptoms, the appropriate therapy is short-acting bronchodilators, as long as your FEV1 is more than 50 percent of predicted.

And so technically, somebody who's FEV1 is 55 percent of predicted and who's not symptomatic, this measure would ask for overtreatment based on -- this is based on gold, which is what we're all using to decide this measure.

So my sense would be this measure would be better if it was for patients with FEV1 less than 50 percent of predicted. Because then all of those patients, according to gold, should be on a long-acting beta agonist or a muscarinic agent.

MS. FRECHETTE: It can be, yeah. The numerator details are below this question, S-6. Specify the patient has to have symptoms, describes what those symptoms are. It also states that a patient has to have FEV, I believe it's less than 60 percent. So that's all in the details of the numerator, which I believe the choice --

DR. PATEL: And somebody, I think that's what I was trying to say, is that our

denominator currently has an FEV1 that's less than 60 percent, and has to have symptoms. It's not an or; it's an and. Would the committee feel comfortable with that or is the committee recommending that we drop the FEV1 to less than 50 percent and have symptoms?

DR. NISHIMI: Well you can't -- just this is Robin Nishimi from NQF interjecting. You can't just sort of change the specs on the fly here without adequate testing. So I just want to point that out.

DR. PATEL: I will --

DR. NISHIMI: If your concern --

(Off microphone comment.)

CO-CHAIR BRATZLER: I just want to make sure, because I don't know the GOLD guide. I haven't read GOLD in a while. So is it okay with at 60 percent with symptoms? Are you comfortable with that in GOLD?

DR. DiGIOVINE: I'm comfortable with that.

CO-CHAIR BRATZLER: Okay.

DR. PATEL: Yes, and that's consistent with GOLD guidelines. It's less than 60 percent and symptoms, would indicate long-acting inhaled bronchodilators as a currently prescribed therapy.

CO-CHAIR BRATZLER: Okay, yeah.

DR. DiGIOVINE: I would agree. I think I didn't read it as closely. I think saying, and, I'm fine with that.

CO-CHAIR BRATZLER: Okay, all right.

DR. NISHIMI: And I just want to confirm. So then given that there is -- because we are discussing gap and you're comfortable now with the specifications, are you comfortable that there's a gap, because that's what we're on?

DR. DiGIOVINE: Did you bring any evidence of a gap? I guess that's what I'm trying to understand. So there's no evidence of a gap?

DR. NISHIMI: It's a new construct.

DR. PATEL: We don't have any evidence of gaps. There are -- there has -- the committee

members do have experience just based on abstracts and so forth.

But there was -- the committee felt comfortable with the knowledge that they -- everyone perceived that there remains a gap when long-acting bronchodilator indicator is a bind. So but we actually don't have any published evidence of that.

DR. O'BRIEN: I mean just mathematically, if all bronchodilators are under-performing, restricting that there's going to be a lower rate by definition. It can't exceed the current performance rate for all bronchodilators, right?

CO-CHAIR BRATZLER: Not much of a gap.

DR. O'BRIEN: I mean the performance gap's tapped out according to the numbers that they have in recent years. But then obviously that's -- it's not apples and apples with the numerator.

CO-CHAIR BRATZLER: All right. So any other conversation about gap, recognizing that

there is no data at this point. There's a perception in published literature that suggests that the gap is going to be there, and their committee was comfortable with it.

But they can't provide that data at this point. So I think we should go ahead and vote on performance gap.

MS. AMIRAULT: Performance gap for measure 0102, 1 being high, 2 moderate, 3 low and 4 insufficient.

(Voting.)

MS. AMIRAULT: Dave, we have not received your vote. Could you please revote? Thank you.

(Voting.)

MS. AMIRAULT: 0 high, 5 moderate, 3 low and 14 insufficient.

MS. BAL: So just policy-wise, since this is a maintenance measure, we can continue to review it, even though it's fallen on gap. We do have a policy called inactive endorsement with reserve status.

So if this measure were to move forward and you would eventually vote on that, it would be marked as so. To be eligible for reserve status though, the measure has to be considered overall an excellent measure, reliable, valid and be considered if it leaves the field, it would hinder it.

DR. NISHIMI: The specifications have changed. So it just cannot be eligible for reserve status. Okay. The numerator has changed, the specifications have changed. So the question is given those changes, do you want to move it into reserve status? Poonam seems to indicate that the policy is that it doesn't matter whether the specifications change. I would disagree with that internally, so we'll air our dirty laundry now.

The question is whether you want to move this into reserve status.

MS. FRECHETTE: Before we go there, could I add one thing? We attempted to find additional information on disparities. We found

a research study that is underway right now and has some pre-publication data. It shows some hints at disparities.

But the other thing it shows is the difference between the prescription of short-term versus long-term bronchodilators. The study is not enormous. It's about 60,000 patients and one hospital system. I don't recall over how many years, but it does show a significant difference between long- and short-acting bronchodilators.

If you're willing to accept some indirect evidence, we'd be glad to submit that or go forward with the reserve status discussion.

MS. GORHAM: So let me just read the policy, so that sheds a little light, just in case we have -- we're a little unclear. So according to NQF policy, if this measure has failed, so it has failed. This must past criteria, which is gap. NQF has the option of granting inactive endorsement with reserve status for measures that meet all other criteria except gap.

So that means that we would vote on the other criteria, and if those criteria pass, then we can have this measure in reserve status. So to continue, this status applies only to highly credible as well as reliable and valid measures that have high levels of performance due to incorporation into standardized patient care processes, and quality improvement access.

Inactive endorsement with reserve status retains these measures in the NQF portfolio, while also communicating to potential users that the measures no longer address high leveraged areas for accountability purposes. The consistent standards approval, which is committee CSAC, knows that the default action should be to remove endorsement unless there is a strong justification to continue endorsement.

So the question would be does the committee wish to continue evaluating this measure for the purpose of -- for possible reserve status, which means we would go through the other criteria and then vote.

CO-CHAIR BRATZLER: And my one question for the NQF process is if the measure is left in reserve status and then over time ATS or ACP or others collect data that shows that there is a substantial gap, then moving it to endorsed status would be easier or --

MS. GORHAM: Then we would go back through the process.

DR. NISHIMI: You would go through the whole process. It has to go through the whole process, correct.

MS. GORHAM: Yes, yes.

DR. NISHIMI: The advantage of being reserved in some cases, not all but in some cases, there are certain statutory requirements for use of measures in certain classes. And so if you remove endorsement, implementers are upset because they may or may not be required to use an endorsed measure.

So that is one of the chief advantages of reserve status. It also sends a signal that the committee and NQF felt this was important to

carry some kind of endorsement imprimatur, because it's an important area, you know, and you don't want the performance to fall off. But the default is that endorsement be removed and I would have argued, as I've indicated, that because the numerator has changed it's actually not eligible. But that's another internal argument we'll have to resolve.

MS. BAL: And I'll just argue that it's the standing committee's decision to make. It's only an option. You do not have to choose to move the measure forward on reserve status. I'm not indicating that's a decision you have to make. That is an option you have. So I just wanted to bring that everyone's attention.

CO-CHAIR BRATZLER: So I think at this point we know we have two options. One is just to stop the conversation, the measure loses endorsement because it doesn't meet the app requirement and then ATS would have to bring it back as a whole new measure in the future.

Versus we go ahead and go through the

process and vote, and as long as it meets all of the other requirements, potentially once NQF determines internally how that works, that it may retain reserved status as a metric. So I think those are our two options at this point. So --

DR. DiGIOVINE: Can I ask one question before --

CO-CHAIR BRATZLER: Sure.

DR. DiGIOVINE: When we move forward, we're going to be asking questions about reliability and validity. Are we going to be asking those questions about the new measure with the new denominator, or the old measure with the old denominator?

And just to -- so to sort of quicken this, if we're going to be asking more questions about the old measure, it just strikes me that we're going to say there's insufficient evidence for any of the things we're going to ask going forward.

DR. NISHIMI: The -- you're to evaluate the measure that's before you. So those are the

specifications that have been modified.

MS. FRECHETTE: If I can ask a question? It seems like if you were to put a measure in reserve status because it's topped out, it would have to be the measure that's topped out, which is the current language.

DR. NISHIMI: And that's why I'm saying that there's a disconnect here, because if you decide to move forward, you're evaluating the measure that's before you. So the reliability, the validity, feasibility and use feasibility. That would be the measure that you're evaluating for reserve status, or you can stop.

DR. DiGIOVINE: But that's not -- but that's not ultimately the data that they provided us.

CO-CHAIR BRATZLER: The other point to ask is whether if this measure which we're evaluating, that's before us with the numerator revision that we've discussed, if this fails on reliability or validity, then it's out; is that correct? That's two strikes and that's it?

DR. NISHIMI: Then it fails.

CO-CHAIR BRATZLER: So then it doesn't make it to reserve status?

MS. BAL: And also a clarification. Right now, you're only determining if you want to consider it for reserve status. You would still at the end of this vote, if you actually wanted to give it the reserve status. So this vote is not -- if you're giving it reserve status, it's just that you want to consider it for that. So I wanted to clarify that.

DR. NISHIMI: So you walk through every single additional criterion, and then you actually also vote separately on reserve status. So it fails one of the mandatory criterion, then it just fails.

CO-CHAIR BRATZLER: Either way, the measure once data's available would have to come back through the whole process, either way.

DR. NISHIMI: Correct.

CO-CHAIR BRATZLER: So I don't know that it makes a big difference. I've read

through the comments from the pre-call, and it sounds like on reliability, validity, usability, feasibility, it is in broad use. It's already a PQRS measure. It's used in the maintenance, you know. It's broadly out there.

I didn't see any enormous questions about the other fields. So I'm going to ask that we --

CO-CHAIR LANG: I was just going to say I was on the call, and I don't think we understood things in the way that you've described them just now on the call. Is that fair for others in our work group?

CO-CHAIR BRATZLER: Well, I'm going to invoke Robert's Rules here.

CO-CHAIR LANG: I'm sorry.

CO-CHAIR BRATZLER: No, that's fine. I'm going to invoke Robert's Rules and just say, you know, is there a motion about what we should do at this point, and then we'll look for a second and we'll vote on it. Yes, a comment?

DR. OHTAKE: I was -- I participated

in work group 1 as well, and we weren't aware of this and just from my reading -- thank you -- just from everything that I'm hearing, it really sounds like to do the right thing with this is to retire it, because it topped out and recognize that this change in the numerator actually makes it a new measure.

And so rather than have it a good measure that's been used to so broadly and its old numerator retired because the performance measure was achieved seems a more graceful way to look at this, and then have this brought back in it with the new numerator as a new measure that can be -- it's a more refined measure that can be seen rather --

You know what I mean? Like I hate to see it become disenfranchised or viewed as a bad measure when it's just a different measure. So just my thoughts.

DR. JIMENEZ: My motion would be to consider this for either retirement and we go through the fail procedure that we just --

CO-CHAIR BRATZLER: It's one or the other. We've got to have -- my motion would request a motion to either we recommend retirement or we go through the process and consider it for reserve status.

DR. DiGIOVINE: I second Patricia's motion.

CO-CHAIR BRATZLER: Which was I think to retire. Okay. We have a motion and a second in the room. Do we have any other discussion about recommending retirement of the measure as a topped out measure.

DR. DiGIOVINE: So I'm sorry. Are you asking to retire the old measure or this -- so that's not in front of us. That's what you're saying. The measure that is not currently in front of us.

CO-CHAIR BRATZLER: No. This is the --

DR. DiGIOVINE: With the different denominator.

DR. NISHIMI: This is the new measure with the new numerator.

CO-CHAIR BRATZLER: But it's the same number though. It's the same --

DR. NISHIMI: It's the same number --

CO-CHAIR BRATZLER: The NQF number that's in use today, that you saw the gap data that says it's essentially a topped out measure.

DR. NISHIMI: So the question for --

CO-CHAIR LANG: State the question again.

DR. NISHIMI: Let me ask the developer this. Is the developer willing to stipulate returning to the old numerator, for which you have data here. You have data that it's topped out. You have testing data. Stipulate that you're returning to the old numerator and then the committee can then decide whether to consider the old numerator, in effect the old measure 0091 as originally specified for reserve status.

MS. FRECHETTE: Measure 102. I'm sorry. You said 91.

DR. NISHIMI: Oh, I'm sorry. Well yeah.

MS. FRECHETTE: Here's my response. The short answer is yes, we would consider that. But to this point that was just made, thinking about what would have the biggest impact on quality. We're thinking is to retire current measure 102, which is the old language, which is just generic bronchodilators, and then submit a new measure with partial substantiating data for -- so the numerators would be long-acting bronchodilators.

So if my understanding of the reserve process is, that it 'll allow us to use the reserve measure for a little bit longer while we're gathering additional information on a new measure.

DR. NISHIMI: Wait. So you're proposing 102.

MS. FRECHETTE: Consider 102 topped out and then --

DR. NISHIMI: The issue we have here before us is that all you've provided is 102 data. So the committee has voted for no gap.

The committee seems a little bit hesitant to march down the trail of placing 102 itself into reserve status. So I'm asking you if you want to withdraw this measure as 102 with the new numerator?

MS. FRECHETTE: And put the original one in reserve status; is that what you're saying? I'm sorry. I don't understand your question.

DR. NISHIMI: 102 now has a new numerator, and it's failed. Are you willing to consider a measure that is 102 as originally specified?

MS. BAL: So going back to the --

DR. NISHIMI: Yes.

MS. BAL: So you would no longer have the new numerator. You would go back to the old one, and that would be the measure you're bringing forward, the old measure with the old numerator, not having the detail that you have.

DR. NISHIMI: All the data you have here are speaking to that. That's the problem.

DR. O'BRIEN: So to clarify are they, or is the reliability actually with the new numerator?

MS. FRECHETTE: That is with the old numerator.

DR. O'BRIEN: Thank you.

MS. FRECHETTE: Oh, any inhaled bronchodilator.

DR. O'BRIEN: Thank you.

DR. NISHIMI: So what does the developer want to do?

MS. FRECHETTE: It sounds like what would help most with basically COPD quality is to withdraw the changed measure, go back to the original language and vote on if it's reserved or topped out or retired, whatever your decision is.

DR. NISHIMI: Because let's understand here. Your reliability testing you updated, but it was still with the old numerator. You have no new reliability testing with the new numerator?

MS. FRECHETTE: Correct.

DR. NISHIMI: Does everyone sort of

understand the baseline fact? Okay. So given that, the question to the committee is do you want to go with topped out? So we would -- we would reverse, you know, and move forward.

CO-CHAIR LANG: Excuse me. So then we would need to revote on performance gap?

DR. NISHIMI: Yes.

CO-CHAIR LANG: Meaning that we don't need to vote on the evidence. But we do need to revote on performance gap.

DR. NISHIMI: Right, but on performance gap.

CO-CHAIR LANG: Because it's a different issue now?

DR. NISHIMI: It's a different issue.

DR. O'BRIEN: Unless we go revisiting the evidence about any bronchodilator versus long-acting being better than short-acting. So the reason as I understand it of the change of the numerator is because if new evidence suggesting that long-acting is better than short-acting.

But now if we're going back to the previous specification saying they're all equal. But I'm hearing us say that that's not what the current evidence supports. I think this measure's going down as the old measure. I don't think it's going to pass this committee.

DR. DORMAN: Well yeah. I mean I would even extend that to say it appears from the conversation that the old measure would go down as retired, and the new measure would just go down because there's no data and we would end up voting another insufficient, which would negate moving forward, because then there's be two insufficients.

So that the -- it seems like the only viable option is to agree to make the numerator go back, make 102 be what it was and then deal with the motion for retirement. Am I --

DR. NISHIMI: So is the developer willing to go back and have the committee now walk through the original?

MS. FRECHETTE: My response is yes.

Bela, do you have any objection to that?

DR. PATEL: No, I agree.

CO-CHAIR BRATZLER: So hearing that, that means essentially start the measure over with evidence.

CO-CHAIR LANG: Right. So we're back to square one. This is a process measure.

(Laughter.)

CO-CHAIR LANG: And the numerator then is patients, patients who were prescribed an inhaled bronchodilator, the denominator being -- I'm assuming it's the same denominator, all patients age 18 and older with a diagnosis of COPD based on spirometric criteria, stipulated here as FEV1 to FVC less than 70 percent, or FEV1 less than 60 percent predicted, with symptoms, dyspnea, cough, sputum and/or wheezing.

And the -- as mentioned previously, updated evidence was provided based on clinical practice guidelines, GOLD, et cetera and I don't know that I in the interest of time I need to go through all of this.

And the issue has been raised about the numerator including both long-acting and short-acting bronchodilators, for which my understanding is there are different indications for these. So the numerator suffers from imprecision.

CO-CHAIR BRATZLER: Any other questions about evidence? We're basically considering the original measure, any short or long-acting bronchodilators. Patients with COPD, denominator FEV1 less than 60 percent and symptoms. Any other discussion about evidence? Richard.

DR. MURRAY: Can I just ask a point of clarification? There was a motion and a second to retire it, and now we're going backwards. So what happened to the vote on --

DR. NISHIMI: Go forward to the retirement.

DR. MURRAY: Does retiring this measure risk any patient harm, by withdrawing that measurement?

CO-CHAIR BRATZLER: Well, I mean I think -- as we've highlighted a number of times today, the reporting of this measure is largely voluntarily. It's people that decide to participate in maintenance and certification and use this measure. It's people who choose this PQRS measure from the list of 300 that you can report.

So to me, the fact that the numbers are high simply tells me the numbers are high for those that have chosen to collect and report this metric. It doesn't necessarily reflect national performance rates, but it's all the data that we have.

So I don't think there's any potential patient harm, and I suspect that there's still a gap in performance. Yes.

DR. SCHINDLER: This is just a point of clarification for the staff. So once it's retired, can it come out of retirement?

DR. NISHIMI: Yes. It's not retired. It's endorsed with reserve status. So it's not

retired.

DR. SCHINDLER: So I guess --

DR. NISHIMI: There was actually a lot of -- there was a lot of thought in not calling it a retirement. So it's endorsed with reserve status and it can come out of reserve status.

DR. SCHINDLER: So the whole notion that it's going to fade off into the sunset gracefully and not come back is a little bit untrue, because the one thing I am concerned about is if it's still able to come back, that there are new data that said maybe this isn't the best measure.

DR. NISHIMI: It's still endorsed. It's just put into what's called reserve status, to indicate that it's topped out.

DR. SCHINDLER: Okay.

MS. BAL: So just for clarification, when you mean retired, I just want to make sure you don't mean if you do not recommend the measure. If you do not recommend this measure to move forward, it does not mean it's gone forever.

It's just like any other measure where it can come back if another project comes up.

CO-CHAIR BRATZLER: All right. Let's keep the conversation moving.

DR. OHTAKE: I just want to confirm with the developer that the denominator did not change, because I had a sense in the conversation that the symptoms may be new and the FEV like the -- that the GOLD criteria in the denominator and the addition of symptoms may be new with this particular application, compared to the old 102 that we're talking about. Can you clarify that?

MS. FRECHETTE: The denominator did not change. I'm sorry. The denominator did change slightly. It's just to basically correct an error in the past, basically a typo.

DR. OHTAKE: Okay. So the denominator in previous -- like the denominator of the previously endorsed measure had the same FEV1, FVC and FEV1 and symptoms that was all in the currently endorsed one?

MS. FRECHETTE: In 2013, there's a

transcription error. So yeah.

DR. OHTAKE: Okay. So it was just a --

MS. FRECHETTE: I'm sorry?

DR. OHTAKE: It was just a typo then, but all the components were there? Okay.

MS. FRECHETTE: Correct.

DR. OHTAKE: I just wasn't clear on the symptoms conversation you had.

MS. FRECHETTE: Yeah, yeah.

CO-CHAIR BRATZLER: Todd.

DR. DORMAN: I don't know what rules we follow, Robert's or others', but I'm still stuck that based upon my understanding, once there is a motion now we can say that the motion's not possible and ask them to remove that motion.

But we either have to vote on that motion or the person who made it has to withdraw it. I don't think we can proceed with a motion on the table.

CO-CHAIR BRATZLER: So I'm going to

suggest -- I'm going to ask that the person that made the motion withdraw it only because we had this point of clarification that we're re-evaluating the entire old measure. So is that acceptable?

DR. OHTAKE: Yeah, that's acceptable.

CO-CHAIR BRATZLER: All right.

DR. O'BRIEN: I concur.

CO-CHAIR BRATZLER: All right, very good. So did you have anything else Todd? Any other comments about evidence for the original measure? Okay. So we have to vote on evidence for the original measure.

MS. AMIRAULT: Okay. Evidence for 0102, 1 being high, 2 moderate, 3 low and 4 insufficient.

(Voting.)

MS. BAL: Would everyone please revote? We're missing a vote.

(Voting.)

MS. AMIRAULT: Okay. 3 for high, 11 for moderate, 6 low and 2 insufficient.

MS. BAL: It passes.

CO-CHAIR BRATZLER: Okay. Discussion of gap, which is where we got to before.

CO-CHAIR LANG: Yes. Briefly, recent data indicate that the gap has narrowed considerably, such that the gap is four percent or less from 2012 to 2014. The gap is very narrow.

DR. DiGIOVINE: Can I ask now? So if we wanted to place this measure in reserve status, what vote would be appropriate to signify that desire?

DR. NISHIMI: Low.

CO-CHAIR BRATZLER: I'd like to ask the developer. So now we are talking about data that you do have, and I know what the national gap is. It's that four percent or so. But what's the range? I mean, is there variability?

MS. FRECHETTE: I received some variability information after we submitted the application, and I just sent it to the staff if you want to see it. But it's a whole report, and

every single report's data is based on different types of PQRS reporting programs. So we have a whole table of data. How would you like to deal with it?

DR. NISHIMI: Well, what's the range? That was the question.

MS. FRECHETTE: Most of range is about 95 percent.

CO-CHAIR BRATZLER: Okay. So we're on variability. Okay. So what we're voting on is the old measure, the data that you've seen before you, the actual gap and does that gap justify continuation of the measure. Is there high -- I'm trying to think of the best way to word this. I can't read that far.

Is there enough variation to say that the gap is a high vote versus low? I don't think it's --

MALE SPEAKER: High, moderate, low.

CO-CHAIR BRATZLER: High, moderate or low. I think we have good data so --

MS. BAL: Okay. Voting is now open

for 0102, performance gap. The options are 1 high, 2 moderate, 3 low, 4 insufficient.

(Voting.)

MS. BAL: All right. The results are 1 high, 0 moderate, 20 low, 1 insufficient, and we would -- this measure would fall on performance gap, but it can be considered for reserve status if you would like.

CO-CHAIR BRATZLER: So as David just highlighted, we're exactly where we were before. So we have to decide. Are we going to go ahead through the rest of the criteria and potentially put the measure in reserve status, or just stop at this point and say it loses endorsement?

Okay, do a hand vote. So all of those who think we should move forward and consider the rest of the criteria to keep this measure in reserve status raise your hand?

(Show of hands.)

CO-CHAIR BRATZLER: Okay. So the majority. So we're going to go on --

MS. BAL: That's 12.

CO-CHAIR BRATZLER: 12 out of 22. So we're going to go through the rest of the criteria. So next is reliability.

DR. BAULDOFF: Yes, okay. Reliability was actually conducted on -- yeah. Was actually -- was actually conducted on the original data for this original measure. So we're in good shape for reliability and validity.

Mathematica Policy Research conducted the reliability analysis on more than 11 million Medicare patients. Reliability came out at .85. According to the reliability algorithm, it comes out as a high.

CO-CHAIR BRATZLER: Any other questions or comments about reliability? Okay. Hearing none, let's go ahead and vote.

MS. AMIRAULT: Okay. Voting is open for 0102, reliability. The options are 1 high, 2 moderate, 3 low, 4 insufficient.

(Voting.)

MS. BAL: Can everyone point again?

(Voting.)

MS. AMIRAULT: Okay. So for Reliability for 0102, we have 11 high, 9 moderate, 1 low and 0 insufficient, and we can move forward to Validity.

CO-CHAIR BRATZLER: And validity.

DR. BAULDOFF: Validity was conducted using a face validity using an expert panel through ATS that reported an 88.9 percent agreement, and the validity algorithm was reported as moderate. Other questions?

CO-CHAIR BRATZLER: Todd, did you have any -- did you have a comment Todd, or your name is standing up -- okay. Any other comments about validity?

(No response.)

CO-CHAIR BRATZLER: All right. Let's go ahead and vote.

MS. AMIRAULT: Okay. So validity for 0102, 1 being high, 2 moderate, 3 low and 4 insufficient.

(Voting.)

MS. AMIRAULT: Okay. So being 1 high,

2 moderate, 3 low and 4 insufficient.

(Voting.)

MS. GORHAM: Susan, can you vote again please?

(Voting.)

DR. NISHIMI: Can everyone vote again?

(Voting.)

MS. AMIRAULT: Okay. For results 4 high, 16 moderate, 2 low and 0 insufficient. Based on the percentage, we can move on.

CO-CHAIR BRATZLER: Feasibility.

DR. BAULDOFF: The data is collected in routine care and it's available electronically. There were no issues with feasibility.

CO-CHAIR BRATZLER: Any other comments or questions about feasibility? Seeing none, we'll go ahead and vote.

MS. AMIRAULT: So for feasibility for 0102, 1 high, 2 moderate, 3 low and 4 insufficient. Again, feasibility 0102.

(Voting.)

MS. AMIRAULT: So 18 high, 4 moderate, 0 low and 0 insufficient, and the percentage we can move forward.

CO-CHAIR BRATZLER: And usability and use.

DR. BAULDOFF: Usability, the data is publicly reported. There were no implementation issues noted, and this was all based on the original measure.

CO-CHAIR BRATZLER: Any discussion regarding usability or use?

(No response.)

CO-CHAIR BRATZLER: Okay. We'll go ahead and vote on usability.

MS. AMIRAULT: So usability and use for 0102, 1 being high, 2 moderate, 3 low and 4 insufficient.

(Voting.)

MS. AMIRAULT: So we've got 10 high, 11 moderate, 1 low and 0 insufficient.

CO-CHAIR BRATZLER: So our last vote now will be about, since we've already said it

didn't meet the gap requirements, about whether it should be -- it's not sustainability, but it's -- should it be -- meet the NQF criteria for endorsed but in reserve status. So, again the vote is yes or no here for reserve status.

MS. AMIRAULT: Okay. Begin voting on potential reserve use, 1 being yes and 2 no.

MS. BAL: And just for clarification if you do vote no and it goes down, the measure goes down.

(Voting.)

MS. AMIRAULT: Okay. So for results, 16 yes and 6 no. 73 percent yes.

DR. NISHIMI: So I just want to make one correction for the record. There were actually 13 votes on favor of continuing and that major change was you went from 14 voting for an insufficient gap, and that turned into the 20 lows, and that's why you were permitted to continue.

CO-CHAIR BRATZLER: Okay, very good. Well, we got through that one. So the next one,

gosh I sat in on the call the other day. This next one should be a breeze, right? 0513, Thorax CT: Use of Contrast Material.

This is one of the CMS measures supported by the Lewin Group. We have our discussants here, Colleen and Charlie. If you would introduce the measure?

DR. BRUETMAN: Good afternoon. Thank you for the opportunity to present NQF No. 0513, which is CT Thorax: Use of Contrast Material. These are studies with and without contrast. There's clear evidence in guidelines that combined studies are usually not appropriate, and this is a measure that is maintained by CMS measure maintained by Lewin and Yale and has been in place and endorsed since 2008 by NQF.

Over time, it has really produced significant effects showing improvement reducing its appropriate use by more than 50 percent reduction. This measure really looks at the ability to reduce overuse of combined studies of the thorax.

Really this measure is just -- the description is calculate the percentage of thorax studies that are performed with and without contrast out of all thorax studies performed, those with and without and with those with and without.

The measures is calculated based on one year window of claims data and it's, as clearly noted, a claims-based measure. So we appreciate the opportunity to address any questions for the committee.

CO-CHAIR BRATZLER: Good thank you. So the discussants were Kim and Ella.

DR. KAZEROONI: Hopefully this is more straightforward than the last one. This should be fairly straightforward. The evidence is fairly clear. There's 36 American College of Radiology Appropriateness criteria that have change from using with contrast or without contrast, away from a combination examination over the last several years.

There's some supporting evidence from

the National Collaborating Center for Cancer and I guess we're just doing the evidence first, yes. So there's extensive evidence to support this measure.

DR. ELLIOTT: Nothing to add.

CO-CHAIR BRATZLER: All right, very good. Any questions about the evidence supporting this measure? Yes, Bill.

DR. GLOMB: If I can just ask the developer to explain. It says the developer chose not to risk adjust or stratify the measure and we had a semantic issue earlier today with risk adjustment and stratification. Can you explain what you all meant by the fact that you didn't risk adjust or stratify and why and perhaps give an example.

CO-CHAIR BRATZLER: So I'm going to ask we hold that until we get to the reliability and validity conversation.

DR. GLOMB: Okay, I'm sorry.

CO-CHAIR BRATZLER: But we'll get to that for sure, Bill. So just on the evidence.

Any other discussion about the evidence? Let's go.

MS. AMIRAULT: So for evidence for measure 0513, 1 being high, 2 moderate, 3 low or 4 insufficient. Again, evidence for 0513.

(Voting.)

MS. AMIRAULT: We have 18 high, 3 moderate, 0 low and 0 insufficient, and the percentage we can move along.

CO-CHAIR BRATZLER: Okay.

DR. KAZEROONI: With respect to the performance gap, as was already described, since the adoption of this measure there has been an improvement in performance. You could see from the data that was prepared in 2001, there was nearly seven percent of all chest EG exams or combined exams down to 3.3 percent in 2015.

But they're considered -- there continues to be wide variation and in the 2015 reporting period, the performance rates varied from zero percent up to 46.5 percent. So there's considerable variation and a gap to improve the

performance.

DR. ELLIOTT: The only thing I would add to that is that were definitely disparities in the size of the facility. 50 beds or less would be more likely to do inappropriate studies. There was also still continued disparities in the data ages 50 to 59 and between black and Hispanic, and a difference between men and women. So there's still gaps that need to be addressed.

CO-CHAIR BRATZLER: Very good. Any other conversation or questions about gaps or disparities? Okay. Let's go ahead and vote.

MS. AMIRAULT: Okay. Performance gap for 0513, 1 being high, 2 moderate, 3 low or 4 insufficient.

(Voting.)

MS. AMIRAULT: If everyone would just give it one more shot?

(Voting.)

MS. AMIRAULT: Actually 20's good. Okay. 13 high, 7 moderate, 0 low and 0 insufficient. On percentages we can move along.

CO-CHAIR BRATZLER: Okay Ella, go ahead with reliability.

DR. KAZEROONI: The measure is very straight forward to measure and therefore both reliable and valid. It is a straight calculation of the number of chest CTs with and without contrast over the total number of exams performed, with and without contrast separately. So it's a very straightforward data-driven metric, easy to calculate and therefore very reliable and valid, purely off of claims data.

CO-CHAIR BRATZLER: And Bill, I wanted to get back to your question about -- this would be the appropriate time to ask about risk.

DR. GLOMB: Thanks. I'm glad -- I didn't know where it fell. Just back to the question about the stratification.

MS. McKIERNAN: Absolutely. So this is Colleen McKiernan from the Lewin Group. So we tested using a logistic regression the impact on inappropriate use of scans with and without contrast, based on patient and facility

characteristics, and we demonstrated that there is a disparity for certain patient and facility subgroups, as was previously described.

We also see a disparity in the literature demonstrating that facility characteristics can impact performance. But we did not adjust for these facility or patient characteristics, because that could potentially mask underlying differences in the quality of care that's provided, and we don't see a reason why the standard of care should differ based on these patient and facility characteristics.

So we thought adjusting for patient or facility characteristics wouldn't mask performance, based on the guidelines that we presented. So consequently we didn't perform any risk adjustment or restratification for those reasons.

DR. KAZEROONI: I would also add with respect to size of facilities, the smaller facilities having the poorest performance could easily be explained by lack of -- the relative

lack of expertise in the smaller number of cases they perform.

The lower percentage of combined studies in women may be related to some consideration of trying to reduce exposure due to breast tissue on chest CT scans. Some of the other ones with respect to age, it's very hard for me to try and figure out a cause and effect reason why age should be related to one group having a lower rate than others, and similarly for ethnicity.

CO-CHAIR BRATZLER: So any other comments or questions about reliability?

DR. O'BRIEN: Do you have any sense of when you talk about the small facilities, the volume of scans done, because in my mind one of the other possible explanations is if you only do two scans. If you do one that actually should get both contrast with and without even though it's not an exclusion, all of the sudden you're an outlier.

MS. McKIERNAN: And so we did control

for minimum case counts. We applied the standards that we applied to when we report the rates to the analysis that we perform. So for that very reason, we didn't want to penalize a facility unfairly if they only had a couple of cases.

DR. DiGIOVINE: Just a quick question. Ella, for this measure, if somebody wanted to some high resolution cuts through a CT with contrast, would that be considered a with and without?

DR. KAZEROONI: There is a definition of the number of cuts with a distance of coverage to count as a second exam. Usually, if you compare chest CTs today compared to chest CTs or any CTs of even ten years ago, the slice thickness nowadays is between one and three millimeters generally relative to the eight to ten millimeters that we saw a decade ago.

So in general, the majority of CTs are done with what would be considered thin sections already. So that's not a practice that is as

readily performed today as it used to be.

CO-CHAIR BRATZLER: Any other conversations? All right. Let's go ahead and vote on reliability.

MS. AMIRAULT: Okay. Reliability for measure 0513, 1 being high, 2 moderate, 3 low and 4 insufficient. Again, reliability for 0513.

(Voting.)

MS. BAL: Can everyone revote please? We're missing one.

(Voting.)

MS. AMIRAULT: 14 high, 7 moderate, 0 low and 0 insufficient, and with the percentages we pass along.

CO-CHAIR BRATZLER: All right, Ella. I think you've already talked about validity. Any other comments about validity or you or Kim, either one?

DR. KAZEROONI: Nothing further.

DR. ELLIOTT: Nothing further.

CO-CHAIR BRATZLER: Does anyone have any questions or comments about validity? Hearing

none, we can go ahead and vote on validity.

MS. AMIRAULT: And for validity, the highest eligible is moderate. So it will be 2 moderate, 3 low or 4 insufficient. Again, we're doing validity for measure 0513.

(Voting.)

MS. AMIRAULT: Okay. The results being 1 high, 20 moderate, 0 low and 0 insufficient, and with the percentage we can move along.

CO-CHAIR BRATZLER: Okay, all right. Feasibility.

DR. KAZEROONI: I think I've probably already covered this. This is very straightforward metric that calculates very feasible taken from claims data. Very feasible.

CO-CHAIR BRATZLER: Any other conversation about Feasibility? Hearing none, let's vote.

MS. AMIRAULT: Feasibility for 0513, 1 high, 2 moderate, 3 low or 4 insufficient.

(Voting.)

MS. AMIRAULT: Okay. 20 high, 1 moderate, 0 low and 0 insufficient, and based on the percentages we'll move forward.

CO-CHAIR BRATZLER: And then lastly use and usability.

DR. KAZEROONI: Since the availability of this metric and the public reporting, we've seen significant improvements, albeit a continued gap and variation in practice. So this is a metric that's easily usable and applied to practice. The data is readily available to all facilities based on the public reporting for use and access.

CO-CHAIR BRATZLER: And Ella, just one question. Has there been any concern about any possible unintended consequences?

DR. KAZEROONI: No. In general, the response is why hasn't it gotten lower faster, and more in the other direction.

CO-CHAIR BRATZLER: Very good.

DR. DiGIOVINE: I'm sort of curious about that too. Given I mean the way the data is

given to us, it seems like there's the -- just about every hospital has this right, and then there's a small number who aren't even close. Is there any sense of like do those hospitals, are they not aware this is publicly reported? Are they not -- I mean it doesn't seem it would be a hard fix to make.

So do you have any sense of why it hasn't had a bigger impact in the hospitals that have these huge rates?

DR. KAZEROONI: It may be lack of awareness. It may be smaller practices that might not have chosen this as a measure to focus on or with a less robust quality portfolio in their facility. I have personally seen small hospitals when bought by larger hospitals and radiology practices incorporated, their rates fall dramatically upon acquisition and implementation of better protocols.

So I've seen smaller practices with this tremendously higher rate compared to the average.

CO-CHAIR BRATZLER: Any other comments or questions about use or usability? If not, let's vote.

MS. AMIRAULT: So usability and use for 0513, high being 1, 2 being moderate, 3 low, 4 insufficient.

(Voting.)

MS. AMIRAULT: So 16 high, 5 moderate, 0 low and 0 insufficient. Based on the percentage we can move along.

CO-CHAIR BRATZLER: And then our last category is overall suitability for endorsement, and I heard very few recommendations to make any changes or any conversation at all. So we'll go ahead and vote on overall suitability.

MS. AMIRAULT: For overall suitability, NRC 0513, 1 being yes, 2 being no.

(Voting.)

MS. AMIRAULT: Okay. 21 yes, 0 no, and with 100 percentage.

CO-CHAIR BRATZLER: Yes.

DR. GLOMB: If I can just make one

last comment before we leave it, and maybe it's a suggestion. I know that for the validity testing we use the claims data. I was curious. Was this all claims data or only paid claims data, given that the point of the measure is patient protection, patient safety as opposed to saving money, did we look at all claims data for these studies?

MS. McKIERNAN: Yeah. So we included 100 percent fee for service data are post-adjudicated. So it's for payment. Obviously, there could be some studies for which there were not post-adjudicated claims, but that's a limitation of the data since we have access. But we can take that point back and see if we're able to evaluate further.

DR. GLOMB: If you've already got the data set, it would be interesting to look at the total denied claims and see what sort of percentages factors are there.

MS. McKIERNAN: Thank you.

CO-CHAIR BRATZLER: All right, very

good. So the last thing on our agenda this afternoon, I know we have a couple of announcements and I'll get to them. But we did have this discussion of related and competing measures to finish up.

So we had two spirometry measures today, one by the American Thoracic Society and the other one by NCQA. NCQA's was a younger age group; American Thoracic Society was specifically looking to see if a spirometry had ever been done. I'm trying to remember what differences there were.

CO-CHAIR LANG: Yeah. I think the major differences were age, 18-40, and then the -- I mean they were both focused on making the diagnosis. But one stipulated a time period for a new diagnosis to perform spirometry and the other was for, you know, basically lifetime, that you know, if the patient went around to different, visiting different physicians for care all of those physicians would be credited in terms of attribution for making diagnosis of

COPD.

So I think -- so I mean my view is that there's clear opportunity to harmonize these measures, and to identify the most desired -- the more desirable features, and then, you know, combine them. I mean both of them are aimed at achieving the same goal in terms of reducing the gap, the performance gap in terms of ordering a spirometry to confirm or rule out the diagnosis of COPD.

DR. NISHIMI: Any other comments for the developers? I saw a lot of heads nodding so -- okay. Then we'll take that to developers and hope that when these come up for maintenance again they will be closer along.

CO-CHAIR BRATZLER: The second one on COPD pharmacotherapies is moot, since one measure is now in reserve status.

MALE SPEAKER: Oh? So that's still in the game.

MS. BAL: Yeah, both are still in. Both measures are still considered in the NQF

portfolio, even if one is endorsed and one is under reserve status. So you would still discuss them for harmonization.

DR. NISHIMI: So if there are features that you like of one versus the other, that would be the --

DR. O'BRIEN: One seems to be more maintenance therapy; the other one is for exacerbations. I think it's useful to have both.

DR. NISHIMI: Okay.

CO-CHAIR BRATZLER: All right. So we have two announcements. I know Robyn you had an announcement you wanted to make.

DR. NISHIMI: I just wanted to make a correction, again for the record on the vote to continue toward the discussion on reserve status. There had been a hand raised, so the vote was 14 to consider, reconsider.

CO-CHAIR BRATZLER: Okay, very good. Then the last thing we wanted to bring up was starting 30 minutes early tomorrow, because I'm told we have some tough measures to talk about

tomorrow. So the suggestion was the continental breakfast would start at 7:30, and the meeting would actually start at 8:00 a.m. tomorrow morning, and then perhaps that will get us out towards the airport on time tomorrow. Is that acceptable to everybody? And I assume --

DR. RILEY: Just a comment.

CO-CHAIR BRATZLER: Yeah, go ahead.

DR. RILEY: Just for anybody who's staying local and planning on taking the subway, there's just an alert issued. It's going to be shut down for at least 29 hours starting at midnight. So if you plan on taking the train tomorrow, don't.

(Laughter.)

DR. NISHIMI: Thank you Crystal.

CO-CHAIR BRATZLER: Be sure that's reflected in the minutes. All right.

DR. NISHIMI: We will end.

(Laughter.)

DR. NISHIMI: Whether you have to have the follow-up call because we don't get through

everything is the open question.

MS. GORHAM: And then a bit of housekeeping. Tomorrow we will choose. Of course this is the standing committee and you all. We ask the standing committee members to serve for two or three years, and so tomorrow by random draw we will decide whether you have a two year or three year term, or you will decide by your draw what terms you will serve.

CO-CHAIR BRATZLER: Any other comments? We're adjourned.

DR. NISHIMI: I'm sorry. Before we adjourn, operator is there anyone who wants to give public comment?

CO-CHAIR BRATZLER: I'm sorry.

OPERATOR: At this time, if you would like to make public comment, press star 1 on your telephone keypad.

(No response.)

OPERATOR: There are no comments at this time.

DR. NISHIMI: Thank you.

MS. GORHAM: Thank you. Have a good evening.

(Whereupon, the above-entitled matter went off the record at 4:57 p.m.)

E

G

C E R T I F I C A T E

This is to certify that the foregoing transcript

In the matter of: Pulmonary and Critical Care Standing Committee Meeting

Before: NQF

Date: 03-15-16

Place: Washington, DC

was duly recorded and accurately transcribed under my direction; further, that said transcript is a true and accurate record of the proceedings.

Neal R Gross

Court Reporter