NATIONAL QUALITY FORUM

+ + + + +

PULMONARY AND CRITICAL CARE ENDORSEMENT
MAINTENANCE STEERING COMMITTEE

+ + + + +

WEDNESDAY
MARCH 21, 2012

+ + + + +

The Steering Committee met at the National Quality Forum, 9th Floor Conference Center, 1030 15th Street, NW, Washington, D.C., at 8:30 a.m., Stephen R. Grossbart and Kevin Weiss, Co-Chairs, presiding.

PRESENT:

STEPHEN R. GROSSBART, PhD, Co-Chair
KEVIN WEISS, MD, MPH, Co-Chair
PETER ALMENOFF, MD, FCCP, Veterans Health Administration
HAYLEY BURGESS, PharmD, BCPP, Hospital Corporation of America
MICHAEL E. CANTINE, BSAST, RRT, CPFT, Morristown Medical Center
RUBIN COHEN, MD, FCCP, Hofstra University School of Medicine
NORMAN H. EDELMAN, MD, American Lung Association
WILLIAM BRENDLE GLOMB, MD, FCCP, FAAP, Texas Health and Human Services Commission
TRUDE A. HAECKER, MD, FAAP, The Children's Hospital of Philadelphia
DIANNE V. JEWELL, PT, DPT, PhD, CCS, The Rehab Intel Network
ELLA KAZEROONI, MD, MS, University of Michigan Health System (by teleconference)

DAVID LANG, MD, Cleveland Clinic

JANET LARSON, PhD, RN, FAAN, University of Michigan School of Nursing
MITCHELL M. LEVY, MD, FCCP, FCCM, Society of Critical Care Medicine
JOHN PELLICONE, MD, FCCP, FACP, Helen Hayes Hospital
DAVID RHEW, MD, Zynx Health Incorporated
CHRISTINE STEARNS, JD, MS, New Jersey Business and Industry Association
CHARLES STEMPLE, DO, MBA, Humana
DAVID C. STOCKWELL, MD, MBA, Children's National Medical Center
CHRISTY WHETSELL, RN, MBA, ACM, West Virginia University Hospitals
DONALD M. YEALY, MD, FACEP, University of Pittsburgh

MEASURE DEVELOPERS:

DAWN ALAYON, National Committee for Quality Assurance
MARK S. ANTMAN, DDS, MBA, American Medical Association
SUSAN ARDAY, Centers for Medicare & Medicaid Services (by teleconference)
KATHERINE AST, MSW, LCSW, American Medical Association
JOHN BOTT, Agency for Healthcare Research and Quality
DALE BRATZLER, DO, MPH, Centers for Medicare & Medicaid Services (by teleconference)
STEPHEN V. CANTRILL, MD, FACEP, American Medical Association
ELVIA CHAVARRIA, MPH, American Medical Association
LINDY CHIN, ActiveHealth (by teleconference)
KERI CHRISTENSEN, American Medical Association
DEBORAH DEITZ, RN, BSN, Centers for Medicare & Medicaid Services (by teleconference)
ELIZABETH DRYE, Centers for Medicare & Medicaid Services (by teleconference)
BRIDGET GULOTTA, MSN, MBA, American Medical

Association
BENJAMIN N. HAMLIN, MPH, National Committee for Quality Assurance (by teleconference)
BRUCE KRIEGER, MD, American Medical Association (by teleconference)
DENISE KRUSENOSKI, MSN, RN, CMSRN, The Joint Commission
RAJESH MAKOL, ActiveHealth (by teleconference)
DAVID NAU, PhD, RPh, CPHQ, Pharmacy Quality Alliance (by teleconference)
SAI NIMMAGADDA, MD, The Joint Commission
DIVYA PAMMANI, National Committee for Quality Assurance
COLLETTE PITZEN, RN, BSN, CPHQ, Minnesota Community Measurement (by teleconference)
MARJORIE RALLINS, DPM, American Medical Association
BOB REHM, MBA, National Committee for Quality Assurance
PATRICK ROMANO, Agency for Healthcare Research and Quality
ELVIRA RYAN, RN, The Joint Commission (by teleconference)
AJAY SHARMA, MD, ActiveHealth (by teleconference)
BANI VIR, MD, ActiveHealth (by teleconference)
ANN E. WATT, MBA, RHIA, The Joint Commission

NQF STAFF:

HELEN BURSTIN, MD, MPH, Senior Vice

President, Performance Measures
HEIDI BOSSLEY, MSN, MBA, Vice President,
Performance Measures
ANN HAMMERSMITH, General Counsel
KATHRYN STREETER
JESSICA WEBER
REVA WINKLER, MD, MPH

ALSO PRESENT:

MAUREEN DAILEY, American Nurses Association
(by teleconference)
SHEILA HEITZIG, American Academy of Allergy,
Asthma & Immunology
MELBA HINOJOSA, Health Services Advisory
Group, Inc. (by teleconference)
DARRYL ROBERTS, American Nurses Association
(by teleconference)

CONTENTS

CONTENTS

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

(8:34 a.m.)

DR. WINKLER: I'm Reva Winkler, Senior Director for Performance Measures here at the National Quality Forum and I'd like to welcome you all to this meeting of the Pulmonary Critical Care Steering Committee.

To the steering committee members, thank you very much for being part of this project. We thank you for the work that you have already done and the work that you are going to do over the next two days.

Katie, do you have your slides up?

MS. STREETER: Yes.

DR. WINKLER: We are intending to establish a phone line for anyone who wants to listen or call in, however we are having some technical difficulties. But I think we can go ahead and get started with some of the introductory things, while they're working out those technical aspects.

So, in terms of introductions, I'd

like to introduce the two co-chairs for this committee, Dr. Kevin Weiss and Dr. Stephen Grossbart, and we're going to be introducing the entire committee in just a moment, asking you both for your introductions and your disclosure statements.

Katie, are we expecting Ann?

MS. STREETER: Yes.

DR. WINKLER: I'm not seeing her. All right. But first I'd like to introduce the NQF staff who are here. You probably are -- know us by name perhaps, if not by face at this point.

First is the Senior Director for Performance Measures here at NQF, Dr. Helen Burstin -- wave to the people. And our Program Manager is Katie Streeter -- you have probably received messages from Katie -- and our Program Analyst is Jessica Weber.

So we may have other staff joining us as they pop in and out. So I think at this point we need to get to know each other with

introductions of the entire group.

We'd like to have you introduce yourself, say a little bit about where you're from, and provide any disclosure statements. Helen, anything else you'd like to add?

DR. BURSTIN: Just in terms of the disclosures -- good morning everybody -- as you go around the room, we know you've already submitted your detailed disclosures. We've got those.

But really the main purpose of today is to disclose anything you think is important in terms of the measures that you're going to be talking about today and tomorrow, and in particular, we'll offer an opportunity after the introductions of the disclosures for any of you to ask questions of any of the other members who talked about their disclosures.

So that's just the introduction and welcome.

MEMBER RHEW: Good morning

everyone. My name is Dave Rhew. I'm an internist, infectious disease physician, health services researcher. I'm from Los Angeles and my disclosure is that I work for Zynx Health and we are an evidence-based clinical physician support company.

MEMBER LANG: Good morning, David Lang, I'm in allergy and immunology at the Cleveland Clinic, and my disclosures are as follows: I have done clinical research for, have served as a consultant for and/or have received honoraria from GlaxoSmithKline, Genentech, Novartis, Merck, Teva, Sanofi-Aventis.

MEMBER GLOMB: Good morning. I'm Brendle Glomb. I'm a pediatric pulmonologist, neonatologist and sleep specialist from Austin, Texas.

I have no current financial disclosures. I am the medical director for Texas Medicaid and the Health and Human Services Commission for the state of Texas.

MEMBER STEARNS: Hi there, Christine Stearns with the New Jersey Business and Industry Association. I have no disclosures. I do policy work for a business trade association.

MEMBER EDELMAN: Good morning. I'm Norman Edelman. I'm an academic pulmonologist based at Stony Brook University on Long Island. I also serve the American Lung Association on a consulting basis as their medical director, and I guess I have no imagination, I have no conflicts to disclose.

MEMBER LARSON: I'm Jan Larson from the University of Michigan, on faculty in the school of nursing, and I do research in pulmonary rehabilitation. And no conflicts.

MEMBER JEWELL: Good morning. My name is Dianne Jewell. I am a physical therapist. I am just down the road in Richmond, Virginia. I was recently full-time faculty at Virginia Commonwealth University in their physical therapy program, but have

discovered the joy of self-employment and now have a consulting business to rehabilitation practices.

I am on the APTA Board of Directors. I don't have any financial disclosures.

MEMBER HAECKER: Hi, I'm Trude Haecker. I'm a pediatrician at the Children's Hospital, Philadelphia. I'm medical director of quality improvement and I'm on the state as the chapter champions for the AAP and I've no financial disclosures.

MEMBER ALMENOFF: I'm Peter Almenoff, pulmonary ICU doc. I'm the director of clinical analytics and reporting in the Department of Veterans Affairs, also on the faculty of the University of Kansas and the University of Missouri, Kansas City, and since I work with the federal government I have no disclosures.

MEMBER STEMPLE: Morning, Chuck Stemple, I'm an ER physician by training, 15

years of managed care, currently with Humana in the clinical policy arena and spend a lot of time in all the Medicare HEDIS quality standards metrics outcomes.

MEMBER PELLICONE: John Pellicone, I'm from Rockland County in the southern tier of New York. I'm a pulmonary critical care physician, still in practice. I'm also the chief medical officer for Helen Hayes Hospital, which is a free-standing, inpatient rehabilitation facility, and I'm also here at the invitation as a board member of the American Association of Cardiac and Pulmonary Rehabilitation. My only disclosures are several non-branded discussions about the COPD and asthma for GlaxoSmithKline and AstraZeneca.

MEMBER WHETSELL: Hi, I'm Chris Whetsell. I'm the director of care management at WVU healthcare in Morgantown, West Virginia. I have no financial disclosures.

MEMBER COHEN: Good morning. I'm

Rubin Cohen. I'm a pulmonolgist working for the North Shore Long Island Jewish Health System in Long Island, New York.

I'm representing the American College of Chest Physicians. I have nothing to disclose.

MEMBER BURGESS: Good morning. I'm Hayley Burgess, director of medication safety and systems innovations with HCA. I have no financial disclosures.

MEMBER LEVY: Good morning I'm Mitchell Levy. I'm an intensivist and chief of pulmonary critical care at Brown University. I represent this side of critical care medicine.

I have no financial disclosures but I am an author on the ventilator-associated events measure.

MEMBER CANTINE: Michael Cantine from Atlantic Health in New Jersey. I'm a respiratory care practitioner. I've worked with Gilead Pharmaceuticals on their Allied

Health advisory board.

CO-CHAIR GROSSBART: Stephen Grossbart. As Chair, I want to remind you to use your microphone.

(Laughter)

Stephen Grossbart, senior vice president, chief quality officer of Catholic Health Partners in Cincinnati, Ohio, and I have no financial disclosures.

CO-CHAIR WEISS: And Kevin Weiss, also Co-Chair with Stephen. I am an internist by training but now the vice president for patient safety and institutional accreditation at the Accreditation Council for Graduate Medical Education.

It's a treat to be here with you all. Stephen and I are looking forward to working with you. As the day goes on, we will get to know each other a little bit better collectively and see how we work together as a group.

Steve and I would request a very

small thing of you all, and that is if you could turn your tent cards at about a 45 degree angle, that would help us and each other so that as we get to know you all -- and we were trying to also recognize the difference between your formal name on the tent card and what you like to refer to yourself as.

So if we didn't get it right on this, just correct us and we'll get through the day better that way.

MEMBER LANG: Yes, I do have a disclosure and that is that I also serve on the board of the American Academy of Allergy, Asthma and Immunology, and am here representing the academy as well.

DR. BURSTIN: So, just to follow up, does anybody have any questions for anybody in terms of the disclosures they have mentioned this morning?

(No response)

DR. BURSTIN: Okay, and lastly,

just one really important thing. Many of you went around and said I represent so-and-so. Actually you are here representing yourselves. You are here for your own expertise. You were nominated by an organization. That's fine. We want to get the breadth of all the various stakeholders involved in this field.

But you are actually here because of your own expertise and you don't represent anybody but yourself today. So thank you.

MEMBER JEWELL: Okay. Since you were kind enough to make note of your effort to attend to the name tent versus the name said, you'll have fun trying to remember that this is pronounced Dee-yon, and not Diane. Complements of my French mother. So I thought I'd give you that heads up now so I don't have to keep correcting you or the staff don't. Thank you.

CO-CHAIR WEISS: Dianne, much appreciated.

DR. WINKLER: All right. Thank you

all very much. I'd like to take this opportunity to just briefly go over some introductory items about the project. Katie, next slide.

Just, we are -- have a fair challenge today and tomorrow, and that is to review 36 measures. There are eight new submissions as well as 28 endorsed measures that are up for maintenance review, in the area of asthma, COPD, pneumonia and critical care.

As you are all very well aware, you were broken into four preliminary workgroups in those four areas to look at a subset of the measures, so you have had an opportunity to look at the measure evaluation criteria and look at the measures.

Those workgroups were intended to take the deepest look at each of the measures and the details. The amount of information for each measure is quite detailed and intense and this was a way of sharing the workload.

Today is the opportunity for those workgroup and lead discussants to share their summary of the information about the measure with the entire group because it is the decisions of the entire group that will determine whether the measure goes forward or not.

The steering committee acts as a proxy for NQF's multi-stakeholder membership. All right? That's why you saw around the table we have a variety of different clinicians, we have a variety of other stakeholders in the room.

So that is fully intentional. The role of the steering committee is to evaluate the measures against the standard criteria. Hopefully, with the work you have done in the -- preliminarily, you are fairly familiar with the criteria.

It's important that we do adhere to the criteria. That criteria has evolved dramatically over the 10 years, 10 or 11 years

of NQF's life, based on feedback from our membership, from users in the field, from developers, from all sorts of folks trying to make them as good as they can be to ensure that the results of -- that the endorsed measures from NQF meet a very high standard.

Ultimately, your job after evaluating the measures, is to make recommendations to the NQF membership for which measures should be endorsed by NQF going forward, all right?

Next one. So I just want to talk briefly about some of the things in the evaluation criteria. I don't want to do a comprehensive review. I think you've had an opportunity.

But there are a couple of things that I'd like to just highlight that I noticed when we were discussing the -- in the workgroups. Katie, next one.

Just recall four major endorsement criteria: importance to measure and report is

not the same as important. All right? So we are asking you to evaluate things not because they are important -- lots of things are important -- but do they meet the criteria as designated in the documentation?

Again, scientific acceptability is not the evidence of the science behind it, it's the reliability and validity of the measure.

Both of those are must-pass criteria. Usability and feasibility I think you are fairly well attuned to. Next.

So, we do have new measures, eight new measures. You can identify those. Their number is greater than 1,000. So if measure 1799, 1800, 1859, those are all new measures.

If the measure is numbered less than 1,000, it is an old measure that is previously endorsed. However, many of the previously-endorsed measures were endorsed at a time when the evaluation criteria was not the same as it is today, and so it is quite

reasonable that measures that have been endorsed for a while, may no longer meet our current criteria, so you shouldn't shy away from the fact that just because it was endorsed before, that it would still meet our current criteria.

And we particularly are looking for data on how that measure is performing now, what's happening, what the current performance is, using that measure.

We want to know about how it's behaving, its reliability and validity in providing us information about quality, how -- what usability issues may have arisen and the same thing with feasibility.

So we are going to be asking you to rate each of those criteria using the generic rating scale of high, moderate, low or insufficient information.

We are going to be asking you to vote on these collectively as a group. You have each been handed a sort of a voting

gizmo, all right?

Each one is numbered. Please make note of your number because we will want you to have the same number tomorrow. We do actually have a record of which one each of you owns.

And so we will be collecting the votes this way. You will be able to see the results as they come up on the screen through the voting software. So on our first measure, we will see how that works. Okay? The next one.

Just to point out the difference between a low rating and one that is insufficient. Essentially low means it did not meet the criteria. Insufficient means there just isn't enough information to know whether it does or not. So there really is a distinction. Next one, Katie.

Importance to measure and report. All three criteria must pass. You will vote on impact, opportunity and evidence

separately, and then the results of those three will be -- we'll look at them on the algorithm and determine whether it passes the whole criteria or not.

Next one. Performance gap. This one is an important one because we discovered in the workgroup that there were several measures whose current performance is very, very high.

We have seen that in maintenance measures, particularly measures that have been out there for a long time, a lot of them that are public reported on a national basis, they are successes.

But at this point in time, performance is very, very high. These are good measures. So, going to the next.

In our cardiovascular project, we ran into this a lot and the committee said, we don't want to remove endorsement from these very good measures, but we know there's really very little opportunity for further

improvement so what can we do?

At the point the board of directors in response approved a designation for endorsed measures called reserve status. And what this means is, yes, they are endorsed but. They are kind of on the shelf, because they simply are not likely to be usable to promote further quality improvement.

But -- so reserve status is something that is -- should be of limited use, because if the measure truly isn't providing any opportunity for improvement, and the concern that there will be a falloff in performance if the measure is no longer used, is really not an issue, then perhaps the measure does not need to remain on our list of endorsed measures.

So, reserve status does exist, and this came up for a couple of measures, on the workgroup call, so this will be in play for us.

Essentially remember that these

measures have got to hit the ball out of the park on all the other criteria. They've got to be really solid. Reliability, validity have got to be good and there isn't a fundamental problem with the measure.

So, you do have that option. So be aware of it. Next one Katie.

Submitted versus existing evidence. We talked a little bit about this in some of the workgroup calls. We really are only asking you to evaluate the information submitted in the forms.

However, how well that's done is highly variable and dependent on the developer, whoever filled out the form. Based on your expertise, you may know additional information.

When you're presenting it, please distinguish from what's presented versus oh by the way, you know, I know that there are -- are other information that would be important that is not included, and we can evaluate it

independently, realizing that anything that is not submitted, is not documented.

Next. We are going to be talking about evidence. Quantity, quality and consistency are the evidence particularly for process measures. For outcome measures, we are not looking for quantity, quality and consistency. We are looking to answer the question, for an outcome measure, are -- is there evidence that there are processes of care that do influence that outcome.

And so, outcome -- looking at the evidence for outcome measures is slightly different than looking at it for process measures, okay. Katie go ahead.

We have given you at your tables each a quick reference to the measure evaluation criteria. Please refer to this if you need to, remembering how to organize the algorithm around evaluating the evidence.

We have given you, on the top of page 2, kind of the decision table on how that

works in terms of whether it passes the evidence criteria.

Okay? So this is meant for your reference. We do have a wide variety of measures that, in terms of some that are very high -- lots and lots of evidence behind it, some less so. Next.

Go on to the next one. Now, there are exceptions to the evidence sub-criterion. There are types of measures that just don't lend themselves to the classic, randomized controlled trials, or even observational trials.

And so the committee will be asked to evaluate what do we know about it, what evidence is presented is reasonable, and so it may fall into this sort of exception to it, particularly if the type of evidence isn't likely to be of the traditional type that you might see in a systematic review or be graded a 1a type level evidence.

So that is allowable as long as we

explain it clearly and the audience is able to understand that you are granting it that exception. Next.

Go on. I think we have kind of gone through that Katie. We don't need to -- scientific acceptability, another must-pass criteria. Just remember we are going to vote on reliability of the measure, and validity of the measure.

Remember that reliability can be tested at the level of a data element, or at the level of the measure's score. Ideally it's tested at both levels.

In order to be rated as high, it must be tested at both levels, and I think you'll find if you look, there are very, very few measures that have been tested at both levels. Generally it's only at one or the other, which is acceptable, but the highest level of rating you can give it is moderate.

Validity a little bit different, because frequently face validity is what's

used. Again, that's going to only allow us to rate it as high as moderate. Okay?

Again, your ratings on reliability and validity we will combine into the algorithm to determine the overall rating for scientific acceptability -- okay, keep going, yes -- according to this algorithm which is in -- also in your Quick Guide. Okay Katie.

Usability, again, is this measure public reported? How is it being used? What do we know about its usefulness? Next.

And feasibility, are there issues about putting this measure in play? If a measure is not being used, the immediate question is why not? Are there feasibility issues? Are there -- is this measure something that is generated through electronic means? Is this a measure that has been retooled for -- as an EHR-based measure? Is it based on claims or other electronic data? These are the real fundamental feasibility issues, as well as susceptibilities to

inaccuracies or unintended consequences. Okay, next.

We will talk about related or competing measures tomorrow after we have had a chance to do a first pass review of all the other measures.

Now, there are a couple of other things that I wanted to point out to you that the Consensus Standards Approval Committee, which is the subcommittee of the Board of Directors, which is sort of the final common pathway for your recommendations and the comments from the membership and public as we reach the end of the process to determine which measures are being endorsed.

CSAC is comprised of a multi-stakeholder panel. They do have a preponderance of consumers and purchasers, as does NQF's Board of Directors.

So we do have a great deal of influence from consumer purchaser perspective at those levels. All right?

They have sort of provided guidance back to measure developers, but it also applies to evaluation of the measures coming forward.

They are not particularly supportive of measures that are met primarily through documentation, the checkbox measure, something that you can just check the box and move on.

They really want to be able to assess the quality of that assessment, that care plan, those advice or instructions. They really -- any impact of patient preferences should be transparent. We want to specify measures for the broadest populations, settings and levels of analysis possible.

Measures of teaching, counseling or advice should be looked at from the patient's perspective. Did they understand, hear and get it?

Exclusions should be supported by the evidence. You should consider the impact

of missing data. Have we pulled out our -- if there's a lot of missing data around an important group of people, and they're not counted in the measure, how does that impact the overall results and our interpretation about quality?

The statistical risk models generally should not include factors related to disparities, so things like race, ethnicity, socioeconomic status should not be risk factors but are encouraged to be stratified when you want to look at disparities.

Adults are defined as 18 years and older and the kind of converse measure where, as you improve, the denominator gets smaller and the numbers change so the interpretation of the measure changes, are very difficult for people to understand as improvement occurs.

So a measure that is going to evolve in that way is less useful for the various audiences going forward. Katie did I

do any more? No. Okay.

We're going to stop with those. Tomorrow we'll pick up on the related versus competing measures when we have that discussion.

But those are just some high points of the issues. I will agree with you that the evaluation criteria has a lot of factors for you to consider. We are asking a great deal of you. This is going to be an intense conversation as we move through the day.

Our job as staff is to help support you as much as possible. If there's any information that you think you need or would like to have to assist, feel free to ask. We can do our best to see what we can get for you.

But are there any questions from anyone? I'd just like to reiterate that this meeting is being recorded and transcribed. The transcription will be posted on our public

website. It's also just the most valuable resource in the world to be able to go back and review your conversation when we are trying to understand how things happened, because you are going to be talking about a lot of detailed information today.

In terms of process, we are going to go through the -- go through each measure. Our measure developers are here. When we get to your measure, I'd ask each of the developers just to, before introducing your measure, to introduce yourselves.

We are going to give them one or two minutes to talk about their measure or groups of measures, as a kickoff. We'll ask the lead discussant to introduce the measure and then we are going to go through discussion, first of the importance criteria, and then the committee will vote on the three sub-criteria there.

We'll move on to reliability and validity, we'll vote on that. If they don't

pass either of those, we'll stop and then we'll go through the rest of the criteria.

Is there any question from anybody on the committee about process, about what we are going to try to accomplish?

(No response)

DR. WINKLER: Then to you. Question? Yes?

MEMBER GLOMB: Sorry, when the presenter presents the individual, not the developer but the presenter, do you want us to walk through essentially the results of our initial analysis?

DR. WINKLER: I think that again we do have some time pressures getting through our agenda, but if we can concisely hit the high points, particularly raise the important, if everybody agreed to hit the criteria, that's a phrase and maybe a why.

If there were disagreements or issues that were raised that are concerns, please raise those.

Okay? Anything else?

CO-CHAIR WEISS: So we are going to go, probably the first couple measures we'll go through a little bit more slowly, just as not only you get your sea legs but Stephen and I are getting our sea legs to help guide us through the process.

What we've talked about with Reva is that we have approximately, on global average, around 15 minutes per measure plus or minus,

And so what we are going to do is we are going to have an eye on the clock for about a 15-minute mark. If we go much beyond that, what we'll do is we'll make you aware that we are going beyond that time, so that we can be mindful of the time process.

So the principal reason for that is, as we all know, with a long list of measures, we could actually steal a little bit of a minute or two or five from each of the early measures and suddenly be without time

and feeling pressured and not giving fair justice to the measures at the end.

So it's just a time management issue. Does that sound like a reasonable approach for all of you? Are you all comfortable with that way of going about it? Are you all with me?

(No audible response)

CO-CHAIR WEISS: Oh good. Good, good. Okay, great. And then the next thing is that I'm going to go through this, and Steve and I haven't done this before, we are going to do a lot of hand-holding with Reva, or Reva is going to do a lot of hand-holding with us as we go through the voting process.

And so we are going to look like we are novices because we are, and so all of you who are in the learning mode, will know we are learning together.

Are we okay with the telephone and everything like that?

DR. WINKLER: Apparently they are

still having some issues but --

CO-CHAIR WEISS: How would you like to handle it for the moment? Do you want to --

DR. WINKLER: I think we can go ahead and introduce -- have -- the first three measures, we could have the developer introduce them.

CO-CHAIR WEISS: Okay. So, we are going to start --

MS. WEBER: We actually have Ben from NCQA. He was planning on calling in. As an alternative we have -- oh.

CO-CHAIR WEISS: No, in terms of a measure developer speaking, because I know David --

MS. WEBER: Yes, he is able to call one of their colleagues here. I don't know if that's acceptable, to have him on speaker phone through --

CO-CHAIR WEISS: Let's try it out and see how it works, if we get a speaker

phone and put it by a mic. So let's just put the speaker phone by the mic, and see how that works.

Welcome Ben.

MR. HAMLIN: I'm sorry. I'm here. What do you want me to do say? I haven't heard anything so far.

MEMBER LANG: Kevin, I assume you want us to introduce our measures.

CO-CHAIR WEISS: Yes, why don't you bring the phone here and then we can talk to Ben and put it by there. This feels almost like low technology. But it's actually very high technology to take a cell phone and --

Ben, you are moving around the room. Hi Ben. You have switched hands. You are in Kevin Weiss's hands for better or for worse, richer or poorer.

I'm going to see if we can try this out here. We are going to put you right by the microphone. What we are asking for is the measure developer, at the beginning of the

measure, to give us a one- to two-minute overview of the measure as -- in any which way you'd like to. So go for it.

MR. HAMLIN: Okay. So we have effectively three measures that are pretty close to a set. That is 0036, 1799 and 1800.

0036 has been in use throughout 2006 in HEDIS. It uses a two-year denominator to identify people with moderate to severe persistent asthma that's been repeatedly validated time and time again.

The two newer measures use the same validated criteria to identify the denominator. 1799 is going to be measured in HEDIS this year, 2012. 1800 is the new one for HEDIS in 2013.

They take a slightly different approach to how they -- the intent of the measure. Medication management for people with asthma, 1799, has different thresholds of 50 percent proportion of days covered and 75 percent proportion of days covered to both

identify organizations that are performing very well in asthma management, but also to try and look for organizations that might be defined, populations that are at risk and might require additional resources or additional attention.

The asthma medication ratio looks at patients' level of controls, proportion of reliever to total medications, such as to be looking at the, you know, how many relievers patients are using overall versus sort of a regular, daily-use controller, and it looks at the medications dispensed only, so we are looking at actual medications that are filled through pharmacy claims.

These are all three -- administrative claims only measures, so they are wholly reliant on claims. A field test was conducted using nine health plans in 2010 with a large number of members and the n for the smallest plan, I think was about three -- two or three thousand members with persistent

asthma. So that's my two-minute elevator speech.

CO-CHAIR WEISS: Okay. Let me then ask, from the committee, if there are any initial questions in response to what the developer has -- what Ben has presented to us.

Okay, so Ben, for the moment, everyone is quiet. We are going to put you down the table and hang out here with us, and we are going to go through the process.

Now I think you should be able to hear us all, but we will find that out shortly. So -- sure. Stay put, Ben. Okay.

Measure 0036. David. So we are going to start with looking at impact, opportunity and evidence and then we will vote on those, and we vote on those as a single or each of those -- each of those separate.

So why don't we just start with evidence. Okay, why don't we do all three together and then we'll vote separately. Okay.

MEMBER LANG: Thank you Kevin. Members of the committee, the 0036 measure focuses on asthma, a high impact condition affecting an estimated 25 million Americans, associated with a cost of more than $20 billion annually.

Asthma continues to be associated with unacceptable rates of morbidity and mortality as members of the committee are well aware.

Performance gaps exist. There are disparities in care and outcomes, and there are opportunities for improvement and there is high quality evidence associated with this.

I believe the measure is -- should be rated highly on the three criteria about which we are voting, and that is what the committee determined in its conference call several weeks ago.

CO-CHAIR WEISS: Short and sweet. Great. So let's have a -- any sort of questions, thoughts, comments, concerns from

the group?

(No response)

CO-CHAIR WEISS: Okay, if none, then we are going to do the first set of votes. Help. Who is going to walk us through this? Jessica, are you going to --

MS. WEBER: Hi, I'll walk you through it.

CO-CHAIR WEISS: Okay.

MS. WEBER: So as you can see, we have the criteria rated one through four, high, moderate, low, insufficient. Make sure you aim your voting device over here towards the voting software, and there will be a clock with the tally so we'll be able to see how many people voted.

Once it's complete and we've gotten all the votes, we'll have a graph of the voting. Make sure you hit the number of your vote, and then send, and if you would like to change it, there should be a little triangle sign at the bottom with an

exclamation point. If you hit the wrong number you can hit that, and then change your vote.

CO-CHAIR WEISS: So we're voting on the impact on -- related to Measure 0036. So let's do that now. High, moderate, low, insufficient.

MS. WEBER: Go ahead.

CO-CHAIR WEISS: Do we know if -- how do we know if our votes are --

MS. WEBER: Let's try it again.

CO-CHAIR WEISS: And there's something we are supposed to point to again? Point to you.

MS. WEBER: Point to Jessica.

CO-CHAIR WEISS: Oh, point to Jessica.

DR. BURSTIN: We'll be able to see the total count at the end, to see if everybody voted, and if it hasn't, we'll ask you to --

CO-CHAIR WEISS: Tell us when to

go.

MS. WEBER: Okay. Go ahead and vote again. All right. Let's try restarting the software and voting again.

CO-CHAIR WEISS: Okay, so we are going to take a 15 second or so --

FEMALE PARTICIPANT: How long does it take you to do it?

CO-CHAIR WEISS: Reboot the computer? Oh, five minutes. Okay, so let's continue on and we'll come back.

DR. WINKLER: We can take the vote -- I think we can take the vote by hand.

CO-CHAIR WEISS: Oh, you want to do it that way?

DR. WINKLER: Sure, why not. How hard could it be?

CO-CHAIR WEISS: I don't know how to do that. I'll -- so we have to vote high, moderate, low or insufficient. How many on -- how many on impact view it as high? Raise your hands.

(Show of hands)

CO-CHAIR WEISS: Okay, let me reverse it so I can make it very quick. How many do not think it's high?

(Show of hands)

CO-CHAIR WEISS: Okay. No, we don't want -- no, this is not a coercion activity. This truly isn't. Would you like to vote moderate or low or something?

(No audible response)

CO-CHAIR WEISS: Very good. So we have one moderate and the rest high. You have to be -- this is not a -- this is -- very good. Okay.

Next we are going to vote on opportunity. And this is the opportunity for improvement, and this has to do with what the measure is, and David, again, your thoughts here were --

(No audible response)

CO-CHAIR WEISS: Good. Okay. All -- these are the same four criteria again. So

all who would say high, raise your hand just to get a feel for it.

(Show of hands)

CO-CHAIR WEISS: Well, let me just do it a little simpler for -- how many would say not high, just so I get a feel.

(Show of hands)

CO-CHAIR WEISS: So there's -- how many of those are moderates?

(Show of hands)

CO-CHAIR WEISS: So, three moderates. All the rest, high. Okay.

And then the third element is the evidence. Is there evidence that this measure is -- good. So let's do it again. I think there will be at least a preponderance on high.

So let's start with one, high. Raise your hands if you think it's high evidence.

(Show of hands)

CO-CHAIR WEISS: Okay. So a

little more uncertainty here. Oh, well, a couple more popped up so let's try again. Raise them up high. High for high.

(Show of hands)

CO-CHAIR WEISS: Okay, how many would say moderate?

(Show of hands)

CO-CHAIR WEISS: One, two, three, four. Five. One, two, three, four, five. That's everybody? Okay. Good. Done.

Well, pushing through technology on to the real, the old-fashioned fallback. Paper would be even worse. Okay. Let's talk about the scientific acceptability, so let's go first about reliability and then validity or you can put them together as part of your discussion. How would you like to go?

MEMBER LANG: Thank you Kevin. Yes. So it's in this realm that I do have some concerns regarding the measure. Let me just frame this from a big picture standpoint, we look at or some individuals around the

table may author systematic reviews. We read guidelines, individuals around the table may serve on guidelines panels.

We identify evidence in the form of practice behavior X, and X is associated with improved patient care outcomes, or from guidelines standpoint, improved population outcomes, in this case, patients with asthma.

What we -- what we want to do is encourage practice behavior X, and discourage practice behaviors Y, Z or A, B, C, which either are not associated with evidence that they lead to improved outcomes, B, may be associated with untoward healthcare outcomes, or about which, C, there are no data which are convincing which show whether outcomes are improved.

So, having framed that, I have concerns regarding the numerator and the denominator of this measure. The denominator of the measure, which is patients aged 5 to 64 with moderate to severe persistent asthma, I

believe there is some lack of precision with regards to moderate to severe persistent asthmatics, and how they are identified.

But the major concern I have with the measure is that the measure categorizes -- seeks to identify use of appropriate medications for people with asthma by identifying the number of members dispensed at least one prescription for a preferred therapy during the measurement year.

And preferred therapies include not only inhaled corticosteroid, about which there are -- there's high-quality evidence supporting exposure to inhaled steroid and improved outcomes, but a number of other medications for asthma, including theophylline and other medications, for which data have not shown that exposure to these medications are associated, or is associated with improved outcomes, moreover, it's -- the number of members dispensed at least one prescription qualifies for fulfilling the

metric.

So from my standpoint, I question whether this metric does what I said at the outset in framing this, in the sense that exposure to, or tracking -- a better way to say it -- tracking medical practice behavior X is associated with improved outcomes.

I don't believe that this measure fulfils that big picture criterion that I have used to approach this. So I question the validity of the metric on that basis.

CO-CHAIR WEISS: Would you like to present the discussion of the group? It sounds like the issue here is validity.

MEMBER LANG: Yes.

CO-CHAIR WEISS: And it's validity of the numerator, not the denominator.

MEMBER LANG: Yes, there is some imprecision with the denominator, but that's not the major issue. The major issue is the numerator.

CO-CHAIR WEISS: And the -- and

you had no concerns with the reliability that we are hearing right now?

MEMBER LANG: No, the matter of concern is validity.

CO-CHAIR WEISS: And would you be willing to reflect the discussion at the group, because we have, at least for the workgroup, it looks like there were three highs, one moderate, which was again very supportive of reliability, and a similar vote on validity.

So, if you can bring in that broader discussion, it would be helpful, if you have recollection of it, of what the group was thinking in response to this.

MEMBER LANG: Yes, I think this issue was raised on the conference call, and I think that there was -- my impression was that a number of members of the committee shared my chagrin.

CO-CHAIR WEISS: So let's now have, if we can, a more open discussion,

starting first, if anyone on the committee would like -- on the workgroup would like to reflect on this issue of reliability, which David is suggesting is solid, but concerns of validity, and what you as individuals might think about that, and then first with the workgroup, and then if we have a more broad discussion on this issue.

MEMBER EDELMAN: I agree.

CO-CHAIR WEISS: Microphone, please. It's all being recorded.

MEMBER EDELMAN: I agree. I think the list of medications has no discrimination. It's too broad.

MR. HAMLIN: Can I come in for a side question?

CO-CHAIR WEISS: In a moment Ben, but let me just make sure that we have all of the reflection we need from the committee first.

MR. HAMLIN: Okay, thanks.

CO-CHAIR WEISS: Any other

thoughts or comments on this, on the concerns for validity?

(No response)

CO-CHAIR WEISS: Okay. So Ben, if you would be so kind, some thoughts on the validity issue?

MR. HAMLIN: Sure. So first I want to address the list of medications. We agreed it's an expansive list, and while the guidelines clearly prefer ICD, you know, recognizing this is a population-based measure, and not every patient is necessarily indicated for ICD.

We want to avoid sort of overriding any critical decision by the provider about what's best for the patient. So the measure list is fairly expansive and it is to try and capture all patients in all scenarios without creating an extremely complex, you know, list of the perfect medications and weighted exceptions for those, in an administrative measure, this is kind of

what we -- this is the list we've been working with.

In terms of the validity issue, the measure denominator has been tested, not only for HEDIS, but also in other, different environments, and has been shown to be quite reliable as a matter of fact, for identifying the appropriate people with persistent asthma.

The people that get in, you know, who might weaken the denominator, it tends to be a very small proportion, running about three to four percent, and it's usually through the ED visit criteria alone --

CO-CHAIR WEISS: So Ben? If I may.

MR. HAMLIN: Yes.

CO-CHAIR WEISS: Because there were no questions about reliability --

MR. HAMLIN: Okay.

CO-CHAIR WEISS: Dr. Lang was focused on the validity as it relates to the numerator and the medicine list, and has there

been any evidence, any studies that have looked at the use of this measure as it relates to other outcomes?

MR. HAMLIN: Not directly to other outcomes, no. The measure has been respecified in other environments to report two rates. They might report an ICD rate and then another rate, but that's as far as we've gone.

CO-CHAIR WEISS: Okay. David, any questions for Ben? Okay.

(Alarm sounds)

CO-CHAIR WEISS: That was the 15-minute mark by the way? So you can see how fast 15 minutes flies. We did have a little bit of a gap there because of the voting process, so we'll just -- this is our first measure.

So I'm just going to reset for another 15 minutes and just so we all get a feel for the time flow here.

So with that in mind, with no

other comments from the workgroup, any comments from the committee at large in response to what you're hearing?

DR. BURSTIN: Ben, this is Helen. Since you mentioned that the measure has at times been used and stratified by ICD versus others, how do the results differ? Do you have any experience in terms of whether different kinds of providers or health plans are going to fall in or out depending on whether you specify it for ICD, which clearly has the strongest evidence, I think?

MR. HAMLIN: Yes, I don't have the detailed data on those stratified rates. I just heard they used the plan and they report back that they really like to stratify it that way. I don't actually have the detailed data. I don't know if I'd be able to get it for you either.

CO-CHAIR WEISS: Yes. Excellent. Well, with that in mind, let's now vote on these electronically, on the scientific

evidence. Sorry. Reliability first. Then we are going to -- reliability, then validity and then we'll look at the scientific -- collectively, I guess.

Okay, so right now, reliability, Measure 0036. Please vote.

(Pause for voting)

CO-CHAIR WEISS: Yes, we're seeing numbers. Numbers are popping up. Oh, there we go. Okay. We have eight highs, nine moderates, one low, and one insufficient evidence. Good.

Okay, now let's go on to the next vote, which is on -- which is on validity. So please vote on validity for Measure 0036.

(Pause for voting)

CO-CHAIR WEISS: So we have 1 high, 11 moderate, 7 low, and 4 insufficient. Okay.

MS. WEBER: We don't have any insufficient.

CO-CHAIR WEISS: Oh, so zero

insufficient. What did I say? Four. Coffee, please?

DR. WINKLER: Okay, and in terms of scientific acceptability we just used the algorithm to -- the majority rated it high or moderate for reliability, or high or moderate for validity, so it passes that criteria.

CO-CHAIR WEISS: Okay, let's move on to the next, which is usability, and then we'll go to feasibility. So let's talk about usability. David?

MEMBER LANG: Yes, the measure has been in effect, as was noted, and information produced by the measure is meaningful, again, with the qualifications that I mentioned previously.

CO-CHAIR WEISS: And so therefore it's -- it's usable? Okay. Members of the workgroup, any thoughts or comments? And you'll see that the votes in the workgroup was three high, one medium. Committee as a whole, any questions?

(No response)

CO-CHAIR WEISS: Then let's vote on that issue of usability.

(Pause for voting)

CO-CHAIR WEISS: Has everybody voted? Peter may have stepped out.

MS. WEBER: We need two more votes, if you want to go ahead and try it again. It won't count your vote twice, if it's already counted.

CO-CHAIR WEISS: Oh, okay. We got what we needed? Okay. There we go. So nine and nine, nine high, nine moderate. Let's continue on with usability to feasibility.

And David.

MEMBER LANG: Thank you Kevin. The measure is feasible. The data are gathered via pharmacy claims, the -- for the numerator. For the denominator, you know, the data are also feasible, gathered based on diagnostic coding.

CO-CHAIR WEISS: Great, okay. Any

comments or -- and your thoughts, therefore, would be?

MEMBER LANG: It's feasible.

CO-CHAIR WEISS: It's feasible. You'd give it a high, moderate --

MEMBER LANG: Moderate or high.

CO-CHAIR WEISS: Okay, and the group was three high, one moderate in the workgroup. Workgroup members, any additional comments?

(No response)

CO-CHAIR WEISS: Okay. And then let's go to the group -- the committee as a whole. Any comments?

(No response)

CO-CHAIR WEISS: Then let's vote on -- oh, we do have a comment from Reva who has an implementation --

DR. WINKLER: Yes, prior to -- when these -- when we launched this project, we posted the list of measures for maintenance and asked for any comments from -- experience

from implementation.

So we do have one comment from AHIP on this measure that says, "We recognize that classification of asthma using administrative data poses challenges, does not allow for tracking and performance by stage of disease, as defined by clinical guidelines.

"As electronic health record data becomes available, it will be important to include clinically-defined asthma stages in ensuring appropriate care by stage.

"Additionally, since a single prescription can ensure compliance, this measure does not track how well asthma is managed for a patient." So, for your consideration.

CO-CHAIR WEISS: Very good. Thanks so much. Any thoughts or comments on what we've heard from the -- Hayley.

MEMBER BURGESS: I'd like to make a comment, based on the discussion of the group.

CO-CHAIR WEISS: If you can get a little closer to your mic it would be great.

MEMBER BURGESS: Sorry. One, I'd like to know if Ben can tell us what the, you know -- how the adherence is currently with the measure, like the percent compliance that we are already seeing with the measure. Can Ben --

CO-CHAIR WEISS: So you want to look at the compliance in terms of use of the measure, or the actual results in the field?

MEMBER BURGESS: The results.

CO-CHAIR WEISS: So Ben, could you just reflect for the group as to what we are seeing in terms of results for the measures in use? If there's some data --

MR. HAMLIN: Yes, so for 0036 we have basically seen, ever since its implementation we have seen a general increase in the rates, where the majority of the rates across the strata, the different product lines, have a relatively high performance,

although there is still a small performance gap.

I had a hard time hearing the question. Was that --

CO-CHAIR WEISS: Yes, you're in line and we're putting up on the screen, I think, the numbers that were submitted to us as well, if you -- on page. So page 13 will be --

DR. BURSTIN: Section 2b5.3 on the submission form, if you want to follow it on your thumb drive.

MEMBER BURGESS: And the reason I asked that question --

CO-CHAIR WEISS: Okay. There you go. So that's overall -- it's called table 3.14, and then we are seeing some of the -- it looks like the mean number was 90.9 percent, or is that -- I can't -- 92.9. My wife, an ophthalmologist who gives me eyeglasses, is going to be upset I can't read that.

Okay, there we go, 92.9. Oh there

we go.

MR. REHM: Just to characterize, there's both commercial rates here and Medicaid.

CO-CHAIR WEISS: So this is commercial rates?

MR. REHM: Yes --

CO-CHAIR WEISS: So 92.9.

MR. REHM: Commercial, basically the 10 percent to the 90th percentile, 89 to 96, and from Medicaid 83 to 93.

CO-CHAIR WEISS: Great.

MR. REHM: So I wanted to make sure that you understood that the Medicaid performance would be an area --

CO-CHAIR WEISS: Is lower and a lot of opportunity for improvement, particularly in the Medicaid population. Is that helpful Hayley?

MEMBER BURGESS: It is helpful. The reason I asked the question is, measures should move us to action, and so if the

measure is that a patient receives one -- one prescription for an asthma controller throughout, you know, this calendar year, does that really tell me how well the patient is doing? Is it giving me something to really work from? Because just saying you've got one prescription, you know, how helpful is that to us? I mean, especially if now we are at compliance in the 90 percent or so. I just question if, you know, this is the right measure for persistent asthma. So I struggle with that a little bit, not that it -- I think it's a bad measure necessarily. I just wonder how that moves us to action because it doesn't tell us that it's appropriate.

CO-CHAIR WEISS: Great and is it Brendle?

MEMBER GLOMB: Brendle, thank you. I wanted to echo that and what David had said earlier. You know, I think there's some -- in the definition of the medication, the appropriate medication, I do think that

there's -- this is a fairly expansive list that many of us around the table might not consider to be controller medications anyway.

So I think that that really makes this a very fuzzy measure and perhaps part of the reason that the compliance rate is so high here.

So I wonder, not that -- maybe the questions we are asking and answering are not approaching what you are talking about. Does it -- if someone has got this diagnosis of moderate to severe asthma, which is somewhat imprecise, and now we've got a list of lots of things that could be precised, is that really getting us -- moving the ball down the field in terms of making an improvement in asthma quality care?

MEMBER STEMPLE: And Kevin, you know, from managed care, I would also reiterate that. I don't know how usable, when you are looking at one script over a year's time, of a broad expanse of medications, the

usability of that data, I don't know how usable it is. We are at a 90 percent for one script in a moderate to severe population which seems total undertreatment.

So when I reflect on the comments and, quote, usability of the data, if we are only requiring one script over a year's time in a moderate to severe population, if we are looking at a quality outcome, that seems a poor quality outcome in a population, one script per 12 months of a broad expanse of medications which we may or may not agree is appropriate.

CO-CHAIR WEISS: It seems to me this is a floor measure, in the sense that at least one script was being written that's not setting a threshold for optimal, by any stretch, is what I'm hearing. Is that --

MEMBER STEMPLE: Yes, and I think, you know, recommendation over time, if we are 90 percent for one, what's an -- what would the pulmonary society say is an acceptable

floor, because I don't think one script for a year, anyone would say in a moderate to severe patient population is an acceptable floor.

So, wondering if there would be an opportunity to move that months of prescription up to a more, quote, acceptable -- what would seem to be a basic floor, because I don't think anyone would even rationally say one is a reasonable floor in a moderate to severe asthmatic population.

But I look to the pulmonologist to maybe reconsider that.

CO-CHAIR WEISS: You know, it's very interesting, in the sense that when this measure came out, it was -- this was considered an extreme advancement, and it may be partly the success of the measure that it's going this way.

I remind ourselves that we are not quite at three sigma here because we are talking about 1 in 10 persons not achieving this in the commercial and up to maybe 2 in 10

or 1 in 5 are not getting even this amount of treatment in a Medicaid population. So it's a great discussion. Peter.

MEMBER ALMENOFF: I do agree with the group but we have to start somewhere, and we still have about 70 percent are not getting a single med, which is actually kind of concerning. So, saying that we are at 90 percent or 85 percent compliance is good, I actually don't think that's very good for something that we've known for a long time, and therapy, we have known for a long time, works.

I mean to me, I think the measure's okay. We just need to eventually develop something better for the future, once we achieve some of our goals. But if we set a measure of perfection, you know, we are never going to get anywhere.

CO-CHAIR WEISS: Yes, and I just wanted -- Reva wanted to highlight for the group that this is also a -- why don't you

speak to it?

DR. WINKLER: This measure has actually been retooled for EHRs and it's part of the meaningful use program.

CO-CHAIR WEISS: So it's, it's now even more embedded in terms of trying to drive this even higher.

MEMBER ALMENOFF: Do I need to disclose that or -- no.

MR. HAMLIN: It's also a CHIPRA core set measure as well.

CO-CHAIR WEISS: It's, say that again?

MR. HAMLIN: It's a CHIPRA core set measure as well.

CO-CHAIR WEISS: Oh, sorry. I was wondering where the sound was coming from, it was in my hand. Sorry. Thanks Ben. You are still in my hand here.

MEMBER JEWELL: Thank you. So I guess this is a question probably for Reva. I have a memory that, from prior panel

participation, that we can make recommendations to the measure developers about things they might do in the future.

And it seems to me that one of the things they might do is really drill into that differentiation that they have already seen with the measure for inhaled corticosteroids versus just the general prescription, so that we can wrap our arms around it a little bit more.

Because I'm guessing, being the non-physician talking here, that there's a greater underuse of the inhaled corticosteroids, but what I heard the measure developer saying is we don't want to overrun clinical decision-making.

So what I don't know is how many patients are likely to be ineligible for those drugs. So we need some data to be able to help better understand that.

CO-CHAIR WEISS: So I'm going to be mindful of time, because I see we are just

about a minute into our second 15 minutes, which is a long space for our first measure but it is our first measure.

So are there any things -- anything that hasn't been said that you'd like to say, as opposed to things that have been said that you want to reinforce?

(No response)

CO-CHAIR WEISS: Okay. Good. So then let's go and vote for the last of the items, which is usability -- feasibility, sorry. We'll get this, right?

(Pause for voting)

CO-CHAIR WEISS: Has everyone voted? Redo yours just in case. You may not have connected. Okay, there we go. So high, 10, moderate, 9. No low and no insufficient.

Very good. Now we go to an overall measure assessment, and it's just a yes/no. Shall we move this on? Now, mind you --

(Alarm sounds)

CO-CHAIR WEISS: That was the timer by the way. So we have officially spent a half an hour on this measure. So again, we're not the final say here. It goes to CSAC, and -- it goes to comment and then the -- oh, back to us and then to CSAC. Thank you.

First measure. So it's a yes/no. Now, mind you I think we have heard it -- we have given to our colleagues and staff that they will let the measure developer know that we do want to see this issue of inhaled corticosteroid more narrowly defined, at least into the future, as an important feature for our consideration.

So let's go for the vote. Yes, no.

(Pause for voting)

CO-CHAIR WEISS: Eighteen -- there you go. You got it. Okay, so 17 yes, 1 no. Great. So this moves on to comment. Coming back to us, and then on to CSAC.

DR. WINKLER: It will go -- you've

passed it this far. It perhaps will come back to you tomorrow, if we talk about related and competing measures, depending on how the evaluation of other similar measures may go forward.

DR. BURSTIN: And just one thought, it might be helpful for NCQA -- and obviously the measure passed -- but I think it would be helpful for the committee to see if you have the data, that inhaled corticosteroids versus all the measures -- all the other meds together, and perhaps even if you have done any sensitivity analyses on the number of prescriptions a year and whether that would make it a better measure as well.

CO-CHAIR WEISS: Great, okay. So Hayley, I think you are up for number two, which is a Measure 1799. It's a new measure. It falls under NCQA. We heard our measure developer describe it initially so I won't ask our measure developer to provide any information now. We'll wait until questions

arise. Okay Ben?

MR. HAMLIN: Sure.

CO-CHAIR WEISS: So we're on the impact and opportunity and evidence.

MEMBER BURGESS: Right, so this is Measure 1799. It is a new measure. It is similar in some ways to 0036, so the committee, our subcommittee, when we went through this first part of impact, we all rated it high, I mean it's very similarly to before, the data hasn't changed. We still believe it is a high impact measure.

So should I stop there for the first part?

CO-CHAIR WEISS: No, let's continue on with opportunity and evidence, if we could.

MEMBER BURGESS: Okay.

CO-CHAIR WEISS: And we'll vote and --

MEMBER BURGESS: So, again, the evidence for, you know, medications in this

space, and this is the same population, if you remember, moderate to severe asthma.

And so here what's different with this measure, and I love the spirit of this measure, so I'll just tell you that. So it's moving from 0036, the previous one, into a space of what they call proportion of days covered.

So now we're getting to the meat of the issue, right? So you know, what extent of time is the patient actually taking the medication, be a proxy of you know, medication database claims.

So that's -- it's a little different in that respect and I really appreciate that because I believe we are getting to the better -- the continuity of care and the appropriateness of care.

It has the same issues, the med list, the drug list is the same for 0036, so it includes inhaled corticosteroids but it also includes short acting beta agonists et

cetera within that drug list so I think that's still a concern.

CO-CHAIR WEISS: Opportunity.

MEMBER BURGESS: So, opportunity. I think there are a couple of things here. What our group -- our subgroup struggled with is this proportion of days covered.

This was a PQA-endorsed phenomenon, if anyone wants to speak to that. I don't know exactly how that translates when we are -- so it's going by number of claims or prescriptions if you are -- and this is maybe a question for Ben or your team -- if it's by claim or prescription, do you have the day's supply, the day's supply for that med?

So insurance companies are pushing towards a 90-day supply, so is that factored in if they get one prescription? Other opportunity, which they do show this in the measure, which I think is really important, that -- because the question is, this is calendar year, right, so the index date is the

first prescription and then days covered throughout the calendar year.

So the question is, if in fourth quarter, they get one prescription, and then it only goes to calendar year, well what does that mean, and that person is still included.

So their data actually does show that the majority of prescriptions are filled in first and second quarter, so it was very low, like five percent I think, in the fourth quarter, maybe 10 percent in the third quarter. So maybe that fleshes out, or maybe that's a place of opportunity, if it's not filled, maybe that's an exclusion if it's not filled within the first three quarters perhaps.

So that's just another thought. And it's similarly tested in the nine health plans.

CO-CHAIR WEISS: Let's -- but that will be coming a bit later --

MEMBER BURGESS: Okay.

CO-CHAIR WEISS: in terms of -- but it suggests that there's an opportunity at least as it's defined --

MEMBER BURGESS: One thing I didn't raise is, you know, this proportion of days covered at 50 and 75 percent, so that's the numerator one and two.

And, you know, the question that I guess the team didn't understand is, is that the right metric, you know, is 50 and 75 percent, is that the right --

CO-CHAIR WEISS: That, again --

MEMBER BURGESS: proportion --

CO-CHAIR WEISS: Hayley, we'll pick that up --

MEMBER BURGESS: Am I moving -- I'm moving ahead.

CO-CHAIR WEISS: Yes, we're moving a little into the reliability/validity issue.

MEMBER BURGESS: Okay.

CO-CHAIR WEISS: But we're looking at opportunity for improvement and what you're

saying is, is that, at least from what they're showing, that there is a place for opportunity, and that the workgroup said -- affirmed that as well.

So let's go and vote that. First of all, from the workgroup, any response to Hayley on issues of impact, opportunity and evidence?

(No response)

CO-CHAIR WEISS: Okay, from the larger group, questions? Brendle and then Peter.

MEMBER GLOMB: Just with regard to evidence, I thought that -- like Hayley, I think that this and actually the next measure also are -- I really like the spirit of where it's going. It's an intriguing measure. It makes sense from a practice standpoint as a treating physician, but I think we are -- if we take the precise definition of the measure, I think we're weak in the evidence area, just as we are in the next measure. I think that

the --

CO-CHAIR WEISS: Weak from what perspective?

MEMBER GLOMB: Weak from -- body of evidence I guess I should say.

CO-CHAIR WEISS: Oh.

MEMBER GLOMB: Weak, in the body of evidence standpoint.

CO-CHAIR WEISS: On the 50/75 percent issue?

MEMBER GLOMB: Yes, on that -- on those cutoffs, yes.

CO-CHAIR WEISS: Okay.

MEMBER GLOMB: Is 50 right, is 25 days right? I don't know.

CO-CHAIR WEISS: Yes, very good.

MEMBER GLOMB: Thanks.

CO-CHAIR WEISS: Peter?

MEMBER ALMENOFF: I'm not sure it's the right time to talk about this, but -- so if you have somebody with asthma and they are put on a corticosteroid, but nothing else,

they'll pass the measure?

CO-CHAIR WEISS: I'll give that to the measure developer. So Ben, the question from Peter was, well, Peter, why don't you just --

MEMBER ALMENOFF: The question is, you know, is there any kind of rescue medication or short term beta2 agonist or something else?

I mean this -- this is for the persistent portion of asthma, but you also need for the rescue piece or for the -- for short term relief.

And so I'm just a little worried when we have a -- such a perfectly-selected measure that, if for example we dumped the first measure we talked about and went just to this, we'd find everyone is on steroid, corticosteroids, which is fine, but we're now not including the other treatments of asthma, sort of just focusing on one medication for one piece of the disease and not the overall

therapeutics.

CO-CHAIR WEISS: Limitation of a --

MEMBER ALMENOFF: Right.

CO-CHAIR WEISS: a specific process of care, as opposed to more comprehensive medication.

MEMBER ALMENOFF: Right, I mean, so on the first one it was too generalized. Now this one is so selective --

CO-CHAIR WEISS: It's the same list. It's just looking at quantitating that as opposed to just -- so it doesn't look at the short acting. It's looking at --

MEMBER ALMENOFF: I understand that, but let's say we just looked at this as a measure and we don't have any other asthma measures, they'll pass this measure and actually not be on the right therapy. They'll be on partial therapy. That's my issue.

CO-CHAIR WEISS: And I would have to say that this is the Achilles heel of any

single process measures, and why I want us to try to create composite measures.

MEMBER ALMENOFF: No, I understand, that's where I was sort of going, is we already have a generalized one, why wouldn't we try to get to a -- a more complete measure as opposed to now we're just sort of doing these partial measures again.

CO-CHAIR WEISS: Okay.

MEMBER ALMENOFF: That's sort of my --

CO-CHAIR WEISS: Good.

MEMBER ALMENOFF: my point.

CO-CHAIR WEISS: Any other thoughts or comments, otherwise let's go now to a vote. First, impact. One through four. Let's do it.

(Pause for voting)

CO-CHAIR WEISS: Thirteen say high, six say moderate. Okay. Let's go to the next one which is impact -- which is performance gap. Thanks. Impact, right? Yes.

Let's vote.

(Pause for voting)

CO-CHAIR WEISS: One more. We got 18 votes, 19. Twelve say high, two say moderate. Okay, let's go to the next and the final of the three.

MS. WEBER: Sorry, seven say moderate.

CO-CHAIR WEISS: I tell you, dyslexia and chairmanship doesn't help. Oh, this is a long day. Let's go. 1c. Yes, it's a great day. 1c. Evidence.

Yes, we have the wrong one down, yes. It's -- no -- it's -- no, evidence is right. Evidence is right. Yes.

(Pause for voting)

CO-CHAIR WEISS: So, 10 say yes, 2 say no, and -- no, sorry. It's the way it's done here, it's confusing me. I apologize. Ten say yes, seven say no, two say insufficient. There we go everybody. I will get this.

That doesn't feel like the right one, right? Yes. We want to do one through four, right? Okay. So, 1c. We want to do evidence. So it is, so it's -- it did pass with -- let's go, next one.

DR. WINKLER: Actually if the vote was 10 yes, 7 no and 3 insufficient, so that's 10-10.

CO-CHAIR WEISS: Let's redo it.

DR. WINKLER: Let's redo it.

CO-CHAIR WEISS: 1c, evidence. One equals yes, two equals no. Three equals insufficient. Let's vote again. Yes, vote again.

MS. WEBER: Actually, there is music, but we don't play it usually. Okay.

CO-CHAIR WEISS: Okay, one yes, two no, three insufficient.

(Pause for voting)

MS. WEBER: We need one additional vote if you want to go ahead and cast it again.

CO-CHAIR WEISS: Everybody cast your vote again. Okay good. So, 16 say yes, 2 say no and 1 says insufficient. Got it. Okay. Next, let's look at reliability and validity, and Hayley, you started talking about those as well, you don't feel the need to repeat yourself, whether you feel comfortable or not repeating yourself. You spoke -- anything else you'd like to say about reliability and validity of the measure?

MEMBER BURGESS: I would like to add one final thought around you know, this percent of the -- you know, possession ratio if you will. If you look at -- well, you guys don't have this -- it's page 13 of the full measure.

CO-CHAIR WEISS: Those who want to, you can go to your thumb drive and it will be on there, or SharePoint if you're logged in that way.

MEMBER BURGESS: So, from the field testing, you know, they broke it out

commercial and Medicaid, so greater than 50 percent, possession ran around 50 percent, greater than 75 percent was around 30 percent, that was in commercial, if you look at Medicaid around 20 percent hits that 75 percent mark.

So you know, really low rates of adherence in this space. But what is good or bad, and is there a benchmark that would come out of that? You know, what is the goal? Is it 100 percent, at the 75 percent? I guess that's the hard part of what we are trying to understand, is do we know enough to say that those markers are still -- are the correct markers?

Though I don't -- I really don't want to say negative things about the measure because I think it's good, I think it's moving in the right direction.

But I think there are some concerns, especially, and still from Ben, would like to hear about the claims, the med

claims.

So if it's a 90-day supply, does that show up, like do they know that in the data?

CO-CHAIR WEISS: So, Ben, that was a question to you. Are they able to actually count, and I know that often they can, the question is in this measure, are they -- is it designed to count the actual number of dispensed days, so that they would pick up a three-month prescription being 90 days?

MR. HAMLIN: Yes, we do actually manage to pick up the multiple of -- multiple canisters if there are multiple or distributed as a 90-day supply. So we actually do count, we count them, and we don't override them. But we do actually -- we were able to count, you know, each day covered from prescription data.

CO-CHAIR WEISS: Great, thanks so much. So then I think what I'm -- if I can recapsulate, reliability not a major set of

concerns from you, Hayley, on validity two concerns, one has to do with how does it treat the individuals who enter into this late in the year, and there's maybe about a five percent at least mis-classification bias that may exist there.

And then, what's the actual threshold and what is 50 percent or 75 mean, is there any evidence that there's a right threshold to be looking at, and those are the two validities.

And for the workgroup, anyone in the workgroup want to comment on what Hayley has said? Does that reflect your thoughts in the workgroup?

MEMBER BURGESS: Can I say one last thing about that? Could we ask Ben, because I think Ben was on the call with us, and we had asked this question of, you know, in the field testing, did they look at outcomes associated with these ratios? Do we know that those that were on 50 percent or 75

percent, did they have a reduced number of ED visits, hospitalizations, et cetera?

Because they had that data, that's what -- you know, that's part of the criteria. That would really be helpful to validate that 50 and 75 percent, that okay, we would believe those are good markers because the outcomes match.

MR. HAMLIN: Sure, I can actually address that.

CO-CHAIR WEISS: Okay. But quickly.

MR. HAMLIN: Okay. So the 50 and 75 percent were selected by a panel much like yourselves, as we actually proposed an initial higher compliance rate much more like MPR of 80 percent, but the panel felt that they really wanted to have two different levels to try and help satisfy the population.

(Alarm sounds)

MR. HAMLIN: We did not conduct --

CO-CHAIR WEISS: That was our 15-

minute mark, just for everyone to know. Okay.

MR. HAMLIN: Oh, right. We did not conduct additional data, but one of the field test sites did go back and look at the ED visits for the population below and above the 50 percent mark, and it did find higher utilization, you got in the in-patient setting for those patients, through the ED, but at below the lower mark and that sort of sub-population was not, you know, very compliant.

They didn't look at the correlation between the 50 and 75 percent so I don't have the difference there, but again, we want -- the respiratory panel, our pulmonary panel felt that they really wanted to see, multiple threshold, it's not a measure that's intended to get up to 100 percent, because you know, we are talking medication compliance.

We did do an additional analysis at the request of -- after the call, looking at the issue of, you know follow-up and the impact on rates.

There's obviously a high correlation with those members who -- especially at the 50 percent rate, who get in in a less than 90 day followup time period.

However, that's less than five percent of the total population, so the overall effect on the rate was almost minimal. More than 70 percent of the members had more than 270 days, which is almost 30,000 members, in a total field test population, had more than 270 days of followup period looking, you know, between the ITSB and the follow -- end of the measure period.

So the bulk of the population was being measured for you know, almost more than half the year.

CO-CHAIR WEISS: I am a little concerned, though, in losing variability and big averages, because that five percent may vary dramatically by health plan and we don't know that data yet, do we?

MR. HAMLIN: We don't actually.

It was -- the window was fairly small. I don't have that chart here.

CO-CHAIR WEISS: Okay, so let's now, with that, just ask -- so what we are hearing is, is that there's been a little bit of testing of the 50 percent threshold, there's been no testing at the 75 percent threshold.

There is a confirmation that about five percent mis-classification may exist, at least on a sampled basis. We don't know what the variability on that is, small health plans, large plan health plans, Medicaid versus commercial, all that kind of stuff.

So that's the information we have. Why don't we go to vote unless there's more questions from the committee. David?

MEMBER LANG: I had a statement, and a general question. The statement is that similar to the previous measure that was discussed, there's concern that I have, and others in the committee expressed in our call

regarding the numerator definition of control with therapy, and that there is inclusion of agents other than inhaled corticosteroid which have not been associated with improved outcomes in patients with asthma.

And then the general question here is that my understanding is that what I just said and what we are discussing, relates to validity, yet it came up during evidence in question 1, and is that an overlap area in terms of a concern that might spill into more than one category?

DR. WINKLER: Actually the question under construct validity is directly, does this measure reflect the underlying evidence, so yes, there is spillover between the two, both in evidence and construct validity.

CO-CHAIR WEISS: Thanks and I just -- a little bit a of a question to me. Are we voting really on one question or two measures? Because there's the two thresholds. Is the 50 percent one measure and then the 75 percent a

second measure? So it's one measure with two parts to it. Okay. I mean I guess one of the question that I would have is if the group felt like they were comfortable with maybe trying the 50 percent, where they may not be interested in the 75 percent. Is there a way of managing that issue if that was to come up?

DR. WINKLER: Essentially we are asking you to evaluate the measure as written, so you are going to be voting on the two. However, there could be a recommendation around developing further data around the 50th percentile and exploring more in the 75th, so you can couch it in terms of a recommendation.

But you are going to have to make your decision based on what's presented to you.

CO-CHAIR WEISS: Okay, so when we get to the issue of validity, then we'll have to link that in. Let's go for the vote, then. So we're on reliability, one through four please.

(Pause for voting)

CO-CHAIR WEISS: Six say high, 12 say moderate, 1 say low and no insufficient. Okay, let's go now to the more discussed issue of validity, again ranking one through four.

(Pause for voting)

CO-CHAIR WEISS: Maybe everyone can vote again just so we can see if we get that 19th vote in.

(Pause for voting)

CO-CHAIR WEISS: There we go. So 1 high, 14 moderate, 4 low and no insufficients. So it passes. Let's go on to usability and to feasibility. And Hayley.

MEMBER BURGESS: So when the group discussed the usability of this, again it was the question of the relationship to the outcome. So the 50 and 75 percent, you know, how does that relate to the outcomes.

But otherwise, felt like the usability of the measure is moving in the right direction, though the concerns we have

mentioned.

Anything the group wants to add?

CO-CHAIR WEISS: I think the only question would be is the public accountability, when you've got the 75 percent uncertainty and I think that was talked about in the workgroup call. Any thoughts or comments on that from the workgroup or -- because this is going to go to public reporting, and other accountability --

(No response)

CO-CHAIR WEISS: Okay, no comments, no questions, then let's go to voting on usability, one through four.

(Pause for voting)

CO-CHAIR WEISS: Four high, 13 moderate, 1 low and 1 insufficient. And finally, usability. Do we have usability as the last one? Feasibility, sorry. Feasibility.

So this is the feasibility. Any comments on feasibility?

MEMBER BURGESS: So with feasibility, it's the same way that they've collected 0036, that measure. So we really didn't have concerns about the feasibility of the collection.

CO-CHAIR WEISS: Great. So any comments from the workgroup? Comments from the group as a whole?

(No response)

CO-CHAIR WEISS: Let's vote.

(Pause for voting)

CO-CHAIR WEISS: Twelve say high, seven moderate, no low and no insufficient. So let's go to the final overall.

CO-CHAIR WEISS: So, suitability for endorsement. Again, this goes out to comment and back to CSAC -- back to us, back to CSAC. Sorry.

(Pause for voting)

CO-CHAIR WEISS: Has everyone voted one or two? Please make sure you vote one or two. Almost there. Good. So, 16 say

yes, 3 say no, and we're done with this measure.

For purposes of quality improvement, we did this measure six minutes faster than the last measure. Let's now go to the next measure. So we're Measure 1800. Brendle.

MEMBER GLOMB: Thank you. Measure 1800 is similar. This is a ratio measurement. This is looking at the percentage of persistent asthmatics, 5 to 64 years of age, who had a ratio of controller medications to total asthma medications, controllers plus relievers, of 0.5 or greater during the measurement year.

Common sense would suggest that they are filling their prescriptions for their controllers, they are saying so well-controlled they're not overly using their relievers, and again, this is a great common sense measurement.

I like the spirit behind it. It's

getting us to where we want to go, patients being controlled and not having to relieve et cetera, appropriate exclusions and there was no risk adjustment of stratification within the measure itself.

I think the committee was very much mindful of the impact of the measure. It's getting us where we have been saying we need to go for so long in controlling the information et cetera, and the rationale is clear.

Like the last measure, some of the quotes were very similar, perhaps these are Hayley's and mine. But concern about evidence within this. So that's the introduction.

CO-CHAIR WEISS: Okay, so other members of the workgroup who'd like to -- oh, this sounds good. It sounds like we've got folks on the call.

MEMBER GLOMB: It's a party line.

(Laughter)

DR. BURSTIN: Hey folks, we

finally got the phones working, so you should be hearing the steering committee discussions now.

CO-CHAIR WEISS: We're on Measure 1800. This is the voice of Kevin Weiss who is co-chairing this. Can you hear us okay on the phone? Oh, they may not be able to respond, right? Okay. Well, welcome. Excellent. Excellent.

So workgroup, thoughts on --

MR. HAMLIN: Kevin, do you want me to redial back in?

CO-CHAIR WEISS: Sit tight. If it's working for you, then please sit tight for this. Is it working for you?

MR. HAMLIN: Okay.

CO-CHAIR WEISS: Good.

MR. HAMLIN: That's fine.

CO-CHAIR WEISS: So, other members of the workgroup want to reflect on what Brendle has said so far as -- I -- on terms of the three elements of impact, opportunity and

evidence?

(No response)

CO-CHAIR WEISS: Good. Okay. Let's go broadly to the workgroup.

MEMBER STEARNS: I just have a quick question. Could you clarify if it is that the -- it's a prescription or whether the prescription was filled.

MEMBER GLOMB: These are claims. So it is a filled prescription.

MEMBER STEARNS: These are claims-based. Okay. Thank you.

CO-CHAIR WEISS: These are dispensed, yes. Okay. David.

MEMBER LANG: Yes, I previously stated way back when introducing the first metric, that exposure to inhaled corticosteroids as we all know has been associated with improved outcomes.

But I think, my concern here is the ratio, that the data are not clear, that this ratio adequately reflects optimal therapy

and leads to improved outcomes.

CO-CHAIR WEISS: So we'll see that as an issue of evidence and validity. You'll come back to this once again in validity. Yes.

MEMBER GLOMB: Kevin, if I may make one more comment. I'm not completely naive to this measure, as Medicaid medical director in Texas. We have been under a federal lawsuit about access to care in the pediatric population for 19 years, and we are under health outcomes measures.

This is actually one of our 10 agreed-upon health outcomes measures with the plaintiffs in this, and what -- my personal experience, our experience in the state of Texas with this measure, is that is most helpful when it is extremely in the negative, i.e. the ratio is very, very low, as opposed to something up there in the middle, which again, takes us all back to the -- I think somebody used the expression sweet spot in our

comments, and I said something very similar, you know, where is the magic cutoff? Is it 0.5, or 0.4, or 0.6, and looking to the scientific validity.

But we -- it has been used extensively since 2007 in the state of Texas.

CO-CHAIR WEISS: So I want to be careful we don't want drift too much into validity, although -- we'll save this and come back to this discussion there.

But what you're saying is the evidence there is a bit fuzzy in terms of the value of this measure, particularly on the up -- as one looks at the higher proportions of the ratio?

MEMBER GLOMB: Yes, and that's where the concerns I think would like.

CO-CHAIR WEISS: Okay, good. So any other comments now from the committee as a whole, thoughts, comments, any --

(No response)

CO-CHAIR WEISS: Very good. Then

let's go and vote. Impact one through four.

(Pause for voting)

CO-CHAIR WEISS: It looks like we've got 19. Is that -- yes. So, 18 say it's high impact. One say it's moderate. Next, we'll go to performance gap. Please vote one through four.

(Pause for voting)

CO-CHAIR WEISS: Okay. Fourteen say a gap of high, five say moderate, no low and no insufficient. Let's go to the third criteria here, which is the evidence.

So this is -- is sufficient evidence, yes is one, two is no, and three is insufficient evidence.

(Pause for voting)

CO-CHAIR WEISS: So 11 say yes, 3 say no and 5 say insufficient. It passes. Let's go to reliability and validity. So let's start with reliability first, if we could Brendle?

MEMBER GLOMB: Yes. Some of the

discussion had to do with the reliability of the definitions, looking at the denominator, at least for asthma medication-dispensing events, where leukotriene modifiers are the sole asthma medication dispensed in that year, issues existing across the measure with regard to the definition of persistent asthmatic, this being overly broad and perhaps imprecise, same with controllers, perhaps overly broad and sometimes unconventional and then how the prescriptions are counted.

And at least on the pediatric side of things, sampling is a consistent part of the process of ongoing care of the patient, particularly in the specialty office, I would imagine the primary care office as well.

So there will be a lot of uncounted medications within this. So the exact ratio, 0.5, aside, I think there's a lot of wiggle room in the definition.

Again, I hate to say too much bad about it, because I think it's a great concept

and it's very practical. It's pushing us further toward our goal. But it is -- you used the word fuzzy earlier, and I think of this as fuzzy in its reliability.

CO-CHAIR WEISS: So I just want to be clear that it's not good or bad, it's just what -- the comments that you speak about in concept, were really A, I mean, where the first vote, where we are looking at, in terms of validity here, is where you are saying that there are certain concerns specifically around this.

Sorry folks on the phone, we are in the middle of what sounds like the entire fire department of greater Washington. Oh, is that the President on the move? Oh.

Yes. Does it look like -- does it look like it's going to go away soon or? We can't even see them. Okay, well let's just punch through it then.

Brendle, in terms of the validity, can you give us another just quick reflection

on the specific concerns on validity that the workgroup and/or you might have had?

Or the lack of validity?

MEMBER GLOMB: Again, I think it falls to the lack of evidence behind this. I know we're beyond that but I can't get away from that, that definition. Maybe somebody else can speak to what the group thought a little bit better than I. I apologize.

CO-CHAIR WEISS: Well, you mentioned earlier the fact that the measure seemed to work better when it's in the negatives than it did in the positives, and do you have a good -- is there some literature or something that helped, or did they provide you with enough information, you will come to all that?

MEMBER GLOMB: No, that's not -- that was our personal experience, or our experience in our state with this exact measure, but it was not -- it was hard to tell where the true cutoff lay.

CO-CHAIR WEISS: Just to the group, ratio measures of course have the problem of a moving numerator and denominator. Right? So you've got this bit of -- you don't know if it's a high ratio because it's the fact they are getting more medicines on the numerator or it's a denominator issue, so there, you can see sort of wild fluctuations there.

From the working group, in addition what Brendle has told us, thoughts or comments?

(NO response)

CO-CHAIR WEISS: And then to the committee as a whole.

DR. EDELMAN: Is the list of controlled medications as broad as we have seen previously?

CO-CHAIR WEISS: Same list.

DR. EDELMAN: Same, yes.

CO-CHAIR WEISS: Trude, you were saying that it was --

MEMBER HAECKER: I'll just echo what Brendle said. Our group was really struggling with this because it's such a broad array of medications, and the evidence around leukotriene inhibitors is not clear. So I think we are all struggling. These three measures all sort of fit into that category.

CO-CHAIR WEISS: So, the general struggle and then applied to this ratio, makes it a little bit more concerning.

MEMBER HAECKER: Yes, makes it even more --

CO-CHAIR WEISS: Okay.

MEMBER HAECKER: Absolutely.

CO-CHAIR WEISS: Good. Any other thoughts or comments? If not, let's go for a vote. This is the reliability, which we have not heard much in controversy of, but let's go for the vote, one through four.

(Pause for voting)

CO-CHAIR WEISS: What's that last vote? So reliability, 11 high, 7 moderate, no

lows and 1 insufficient. Let's now go to validity, where there have been concerns, and so please feel free to vote with your conscience here.

(Pause for voting)

CO-CHAIR WEISS: Almost there, 17. One voted. Vote again everybody, just in case. Okay. So only 1 high, 11 moderate, 4 low and 3 insufficient. So it passes.

Let's go on to the last two criteria, usability and then feasibility. So usability, Brendle?

MEMBER GLOMB: Thank you. Again, looking to meaningful, understandable and useful, I think that this is -- I think again, back to the spirit of the measure, I think it is a very meaningful measure and I believe the committee felt that there was overall moderate evidence toward that.

It was certainly an understandable measure, although the definitions per se are a bit fuzzy, and I think it's probably useful

for intended audiences if we look at both public reporting and, more importantly, quality improvement.

But as to the overall meaningfulness of the exact ratio, again, I think that's where everybody had trouble.

CO-CHAIR WEISS: Well, that's not inconsequential.

MEMBER GLOMB: No, it's not. Yet even the subgroup came up with a predominantly favorable scoring for this.

CO-CHAIR WEISS: Yes.

MEMBER JEWELL: So I guess I have a question for the workgroup. Relative to the -- or maybe the whole group -- relative to the concerns that have been expressed about the medications that are on the list, if I put my Joe Q. Public hat on, I might be able to understand a ratio and I probably could understand controller versus rescue.

But if I don't have the ability to know or understand which medications really

are best in class for either of those two functions, and there are potentially medications on the list that aren't best in class for those two functions, how useful from a public point of view is it?

Is there more advantage just to get the public thinking about it than there is -- not harm, but disadvantage to them being in the dark and not really getting all they could out of it?

CO-CHAIR WEISS: So, if I can summarize that very succinctly, just the very basic question is, how much this ratio helps in public thinking. Okay.

DR. BURSTIN: Just one important note, I mean NQF-endorsed measures are used for a variety of accountability applications, so public reporting to the public is one part of it, but certainly, you know, Christine could talk to purchaser views of this, other views of it.

So there are multiple uses as

well. Point still stands.

MEMBER GLOMB: Just a quick comment. I do think, though, that it -- I think your point about getting the public thinking about that ratio, not so much as a mathematical ratio, but I need to be using the controller frequently, then I will use less of the reliever.

(Alarm sounds)

MEMBER GLOMB: I think that, I think that's a part of it probably because I think our patients see asthma medications in a big bag that they reach into and grab.

CO-CHAIR WEISS: Once again, that was about the 15 minute mark, but we are well into this measure, so other thoughts or comments on the usability and -- what we've heard so far, if I may summarize in a snippet, is that it's -- there's a little but of fuzziness to understanding how it will be used, but that there's a general sense of this is the kind of direction one wants to go.

Is that what I'm -- it doesn't sound enthusiastic but it sounds directional.

MEMBER GLOMB: They're not enthusiastic enough about this, now. Yes, we're very enthusiastic about -- it is advancing the cause and it is perhaps a better measure than some of the other or more outdated measures.

CO-CHAIR WEISS: Okay. Very good. I'm sorry what was that? Yes, fine Ben.

MR. HAMLIN: This measure in particular has been shown to be extremely sensitive in identifying the association with people, it's particularly sensitive in identifying population as far as targeting specific cohorts.

CO-CHAIR WEISS: Okay, what he said was is that this -- at the point five mark, that this has been shown to be effective in its relationship -- directly relationship to ED visits.

So those are -- so Ben, if you

could give me back -- so there's a threshold of 0.5 is the mark?

Do I have that right Ben?

(No response)

CO-CHAIR WEISS: This is not a threshold measure, is it? It's 0.5 okay. Great. Yes. Okay, good. So Hayley is that helpful to you? Good, okay.

Then -- what kind of ED visits? Just a variety of visits or what --? I think it's just an ED visit. Okay? Good.

So let's vote on usability.

(Pause for voting)

CO-CHAIR WEISS: And then let's go to the thinking. While that's accumulating, why don't we think a little bit about the last, which is -- oh, what's -- that quick? Four high, 14 moderate, one low and no insufficient.

Not as enthusiastic here but it sounds like it still passes. Okay. And then feasibility, that should be straightforward.

MEMBER GLOMB: I think feasibility is straightforward. I think it would be even enhanced by electronic data collection, mixing claims versus what's going on in the care setting.

And there -- the only concern I think that the group had, and this had to do with susceptibility -- that foresees susceptibility to inaccuracies and unintended -- more so than unintended consequences, and that we were going to be, because of the broad -- the broad definitions of controller medications and some fuzziness in the diagnosis, that that was some room for these inaccuracies to occur.

But overall, the subgroup was more enthusiastic about this and had fewer concerns.

CO-CHAIR WEISS: Great. Workgroup. Any additional thoughts on what Hayley has said -- not Hayley, what Brindle said?

MEMBER BURGESS: The data source here does list paper records.

CO-CHAIR WEISS: Say that again Hayley.

MEMBER BURGESS: The data sources here, it does list paper records. Can you speak to that a little bit? I don't remember now the conversation --

CO-CHAIR WEISS: Maybe we could have Ben speak to the fact that there's an alternative way of collecting this. Ben, the paper record approach?

MR. HAMLIN: I'm sorry. I can't hear you.

CO-CHAIR WEISS: Okay. Maybe if we can bring this --

DR. BURSTIN: They're going to be fixing it at break. Just repeat the question.

CO-CHAIR WEISS: Okay. So Ben, the question is, is there a paper method, a medical record audit model for this measure?

MR. HAMLIN: Yes, there is, take a

look at the state medical record versus this measure.

CO-CHAIR WEISS: Okay, thanks. Excellent, well thank you. Any other questions, comments? Otherwise let's go to vote for feasibility, one through four, high, moderate, low insufficient.

(Pause for voting)

CO-CHAIR WEISS: Good. Everyone voting one through four. Can everyone revote, just to make sure we are picking up that 19th vote?

So, 13 high, 6 moderate, no low, no insufficient information. Let's go to the final summative vote. Yes, no.

(Pause for voting)

CO-CHAIR WEISS: Sixteen yes, three no, and that completes this measure. This time we did the measure even two minutes faster, so we are slowly getting up to that 15-minute mark. But in the interim, you all deserve a great break, stretch. We have five

minutes. Mainly we are moving the break up because we want to get the phone fixed, but it's also a good time to get a break.

So, how long a break? A 10-minute break. Thank you all.

(Whereupon, the proceedings in the foregoing matter went off the record at 10:22 a.m. and went back on the record at 10:40 a.m.)

CO-CHAIR WEISS: So we have our last person to make us a full complement. Don has made it from Pittsburgh. Do I have that right? So if you could just say a quick hello to the group with your mic on so it's recorded, and also, just a moment about any disclosures, conflict of interest disclosures that you would like to make that may not have been mentioned on your paperwork. Do I have that right? That is included in your paperwork that we haven't seen or have seen. Anyway, just anything you have to say.

MEMBER YEALY: Okay. Thanks very

much. I'm Don Yealy from the University of Pittsburgh. Nice to be here. I apologize for the tardiness.

I don't think there are any conflict of interest disclosures. I'm just working on one NIGMS-funded sepsis trial that falls outside of any of the topics that I was commenting on.

CO-CHAIR WEISS: That sounds great. Well, welcome. I'm Kevin Weiss, and Stephen, do you want to say a quick hello? Oh, let's do that. David Stockwell, you showed up in the middle of the measure process, so why don't you give us a quick hello?

MEMBER STOCKWELL: I did. My apologies. I am a Washingtonian but underestimated the challenge of driving to downtown this morning. It's quite arduous.

So, David Stockwell, I'm a pediatric intensivist here in town at Children's National Medical Center. I am also the executive director of improvement science,

essentially doing quality and safety for our hospital as well, and appreciate the invitation and already enjoying the discussion and the work that's been done to this point. So thank you.

CO-CHAIR WEISS: Sounds wonderful. Welcome on board. Do you want to say a quick hello? Want to say hi just as co-chair or just --

CO-CHAIR GROSSBART: And I just want to introduce myself as the co-chair, Steve Grossbart. Nice to meet you.

CO-CHAIR WEISS: Great. Okay let's continue on then with 0047, which is our first measure from -- today from the AMA PCPI. We have measure developers here, and I think Mark, Mark Antman is going to give us a one- to two-minute overview.

DR. ANTMAN: Yes, thank you. Again, I'm Mark Antman. I'm director of measure development operations for the PCPI which is convened by the AMA.

0047, as you have seen is a measure focused on patients with persistent asthma who are receiving long-term control medications.

Because that measure is very obviously similar to Measure 0036 that you reviewed before, I'll take a moment to just highlight some similarities and differences.

0047 is specified at the clinician level. The persistent asthma population, that is the population of patients with persistent asthma in the denominator of the measure, is defined a little bit differently than in Measure 0036, and I can speak to those differences if desired.

I'll also note that the numerator of our measure does include the alternative long-term control medications, and I'm happy to speak to that as well.

I will -- at the moment though, I will note, I will point out that we did note after the measure had been submitted to NQF

that we unfortunately had some -- a few errors in the list of medications.

We have since corrected those but I have realized in the last day or two that there's still a couple of errors that remain in our medication list, so I'm happy to speak to that when the discussion ensues.

A disharmony that the group may have noted and that I think was highlighted in the workgroup call with Measure 0036 is that the age ranges and exclusions do not match. I'm happy to speak to that as well.

As far as the use of the measure, it's been in the CMS PQRS program since 2007. Our recent testing has demonstrated that the measure is valid and reliable, and finally I'll note that we submitted claim specifications but we also have submitted an electronic measure for them -- for this measure as well.

CO-CHAIR WEISS: Excellent. Any general questions for our measure developer

before we start into the detailed discussion?

If not -- yes, Norman.

DR. EDELMAN: Yes, I notice in the list of controllers, you have a long acting, inhaled beta2 agonist listed between two commas that is listed as monotherapy.

In view of the recent guidelines, shouldn't that be revised so that it includes a combination with inhaled steroid? The way it's listed now it could be used as monotherapy and that would be contrary to current guidelines.

DR. ANTMAN: So that is one of the errors that I referred to. We -- there was a previous version of this measure for which we had the long acting beta2 agonist listed, as well as, and I think -- I'm looking on the screen -- I believe the short acting are listed here as well, and that is one of the errors that we noted.

So I apologize. The long acting and the short acting beta 2 agonists should

not be in that medication list.

CO-CHAIR WEISS: Thanks for bringing that to attention, and Brindle.

MEMBER STEMPLE: Sorry, one similar comment. Was there any thought given to moving the inhaled steroid -- the ICD/LABA combos into the first numerator along with inhaled corticosteroids alone? Was there consideration to given that, and then leaving the others as the alternative controller medications?

DR. ANTMAN: I believe there was some consideration given to that, but because the NHLBI guideline is so clear that ICD are the preferred meds, the workgroup felt that it was more appropriate to state ICD as the preferred, and everything else, including combinations as alternatives.

CO-CHAIR WEISS: David.

MEMBER LANG: Yes, I was going to raise this as I'm going to lead us through this, I assume next. But as long as we are

right on this issue, if you could clarify this, you have three rates as opposed to the other metrics: patient's prescribed inhaled corticosteroid, that's number one; number two is patient's prescribed other alternative long-term controllers; and then three is a total.

Now the focus, at least the concerns I should say, seem to be on number two, which is other alternative long-term controllers, and I mean it's a lot of apples and oranges here in the, you know, in terms of this list.

But when a patient receives one of the three inhaled steroid long acting beta agonist combinations, does that patient also -- is that patient also counted in category one, as receiving an inhaled steroid, or not?

DR. ANTMAN: So I believe the intent is for patients to only be counted in one category or another. And I'll add if I may, Dr. Lang, that the intent of the

workgroup in asking for these three separate rates, was because of the fact that the group noted that it would be of great interest for quality improvement purposes to know how many patients are in fact receiving ICD, how many receiving the alternatives, and what's the total?

So the intent was to tease out that information.

MEMBER LANG: So, just to clarify, the patients, in order for your measure -- well, let me say it a different way. Patients who are in category one or patients who are receiving inhaled corticosteroid, well not monotherapy, but a prescription for an inhaled corticosteroid that is not a prescription for an inhaled corticosteroid combined with a long acting beta agonist. Is that correct?

DR. ANTMAN: Okay, so I realize I think I misunderstood your question in the first place. Give me a moment, if I may, to look at the specifications and I'll be better

able to answer your question.

CO-CHAIR WEISS: We'll come back to that. It sounds like it's a question of validity, principally. Okay, so we'll come back to it specifically there. So you've got a little bit of time. Not a lot, a little bit of time.

Okay, so let's start with impact, opportunity and evidence. I think it would be fine to say impact, to the extent that we've already had that discussion, do you feel like we need to spend more time -- okay. So then let's go to opportunity and evidence, in terms of, David, your thoughts?

MEMBER LANG: Yes, well -- yes, so I think there is opportunity for performance improvement, and I think that the -- again, just to highlight the distinction of this measure compared with the previous measures, is the age group, which is 5 through 50, we discussed that on the conference call. We talked about floating that upwards.

And this -- the denominator, patients with persistent asthma, and again, the issue, the -- I guess this gets to validity but it overlaps with evidence, I've learned in recent discussions, so I'll mention it now, and the issue is the concerns with the numerator definition.

CO-CHAIR WEISS: Okay. So workgroup members, any comments on impact, opportunity, evidence that you'd like to add to what David has said, to the reflection of our study, I mean of our workgroup discussion?

(No response)

CO-CHAIR WEISS: Okay. Committee as a group? Now, you'll note that in the workgroup, there was yeses principally, with the exception of -- that was it. It was principally yeses. Any question from the committee, since the workgroup itself has --

MEMBER GLOMB: Just a quick comment if I can. I, you know, this one head to head with 0036, that we started with,

really --

CO-CHAIR WEISS: Tomorrow's the comparison.

MEMBER GLOMB: Okay. Okay. Sorry I'll get that --

CO-CHAIR WEISS: We'll get to do that.

MEMBER GLOMB: I'll wait until tomorrow.

CO-CHAIR WEISS: With excitement. This will be good. Let's vote. Okay.

(Pause for voting)

CO-CHAIR WEISS: And Don, all you do is you press the number and then send, make sure you press the send after the number, one through four.

Got it. Okay. So, 20 -- well, that was even -- next. Let's do opportunity, performance gap or impact. Performance gap. So this is the opportunity, which is the performance gap, one through four.

What we heard from David was, is

that there was an opportunity -- did you want to talk about the actual number, the proportion of the gap, or is that what you are thinking about?

DR. BURSTIN: I'm sorry. I was just curious if you'd look at the actual performance on PQRS, because you do have 2009 data in here, but it's --

CO-CHAIR WEISS: Did you comment for a moment, Mark, on PQRS, if you have it -- at least if --

DR. ANTMAN: Yes, we did include some PQRS data. We do have a member of our testing team here who can respond to any particular questions about those data.

CO-CHAIR WEISS: Just the overall performance, what it was --

MS. GULOTTA: The gap for 2008 was a little over 46 percent, 46.29 percent.

CO-CHAIR WEISS: Okay. Great. Yes. So let's vote.

(Pause for voting)

CO-CHAIR WEISS: Oh the other thing, Donald, is you need to point it to Jessica when you can. It seems that yours is working, but just in case it --

So, 15 say high, 5 moderate, no low and no insufficient. Next one, which is the evidence. This is a one, yes the evidence is adequate, two is no and three is insufficient evidence.

Okay, so it's one, two, three.

(Pause for voting)

CO-CHAIR WEISS: Almost there. Let's all revote again. Just punch it again. Not change your votes. Just punch it again. This is not Chicago.

There we go. Okay, so 19 yes, one no, and no insufficient. Let's go on to reliability and validity.

MEMBER LANG: So again, some of the issues that have been mentioned previously, regarding validity, in terms of concerns with the numerator definition.

I think I'm wondering whether there's a clarification on the issue of whether patients who receive prescriptions for inhaled steroid, long acting beta agonist combinations are considered in group two. Do you have a clarification on that?

DR. ANTMAN: Yes, I do. Looking at our definitions for the numerator, we do say that the group, group two includes inhaled steroid combinations, so the intent is for anything combined with ICD to be in the second group.

MEMBER LANG: Well, in view of that I would say that I would have some serious concerns regarding the validity of the measure, because the group one would include patients receiving inhaled corticosteroid alone prescriptions, that is not for the three combinations which are frequently prescribed for patients with moderate to severe persistent asthma and that is supported by high quality evidence.

The group two includes patients who are receiving, as has been pointed out previously, agents which are not associated with improved outcomes in the case of long acting beta agonist therapy, could be as monotherapy could be associated with untoward outcomes.

So I have some serious concerns regarding the validity of the measure on that basis.

CO-CHAIR WEISS: Okay, and other members of the workgroup, your thoughts on David's comments?

MEMBER HAECKER: They're valid, excellent points and I think we need to consider those.

CO-CHAIR WEISS: Okay, to the committee as a whole? Thoughts or comments on what you've heard with regards to validity, not so much reliability. Do you have any reliability concerns that you wanted to note?

MEMBER GLOMB: Quick comment

regarding validity. When we translate this to an outcome, is, again, the kind of the time window for the numerator, it's a single controller prescription within the time period, and one questions whether or not that equals control and therefore good outcome.

CO-CHAIR WEISS: I think what we heard of this was in the discussion with NCQA is this measure now enough, or are we moved on? At one time maybe it was enough, kind of feel to it. Okay? Good.

Any other thoughts or comments, otherwise we are going to a vote. Comments, questions?

(No response)

CO-CHAIR WEISS: Okay. So let's vote. Reliability. One through -- oh, Mark.

DR. ANTMAN: If I may, I'd appreciate a chance to comment on Dr. Lang's question.

CO-CHAIR WEISS: Yes.

DR. ANTMAN: So with regard to the

-- to there being some medications in the second list, in the alternative list, that are not necessarily associated with best outcomes, the workgroup was very deliberate about looking at all the medications that are documented in EPR3, in the -- sorry -- in the guideline update that were supported by the guideline as recommendations as alternative therapy for patients with persistent asthma.

I do acknowledge that certainly there are some medications for which the guideline states that there is B level evidence, or I believe, as I recall, I think there are some medications that -- where I think at the very least they state B level evidence.

There are several medications on this list, including cromolyn and nedocromil and the leukotriene modifiers where the evidence is at A level, as it is for ICD.

So it -- we do believe that the list of alternative medications is very

consistent with the most recent guideline update.

CO-CHAIR WEISS: David, is that your understanding of the guidelines as you think about the evidence, because that's --

MEMBER LANG: Yes, I think it depends on -- I actually was going to say, Kevin, I thought we'd get through this in under the -- beat the clock.

CO-CHAIR WEISS: We will, if you can do it in about 15 seconds.

MEMBER LANG: Yes, right. I'm going to try to -- I'm going to try to be brief in my response. You know, I appreciate what you're saying. It depends on the outcome, I guess. You know, there are randomized control trials, so improved outcomes.

I guess what I'm focusing on is improved outcomes from a population standpoint reduced mortality, morbidity, reduced emergency department visits, hospitalizations,

reduced rate of exacerbations over time, in terms of those outcomes, as opposed to, say, you know, spirometric measure over a -- you know, the course of a 12 --

(Alarm sounds)

MEMBER LANG: or 16-week study. There you go. You know, I think the evidence is not as solid in terms of the outcomes I've mentioned for agents such as methylxanthines for instance. Mesostabilizers are not available in terms of cromolyn and nedocromil as you well know --

CO-CHAIR WEISS: David, so, let me just be mindful of time.

MEMBER LANG: Yes, yes, yes. But you have apples and oranges here in terms of antibody inhibitor, and, again, one could fulfil criteria for this metric by long acting beta agonist monotherapy prescribing as well as short acting inhaled beta agonist which is category two.

I mean, that, that -- registers I

the same way as a prescription for one of the inhaled steroid long acting beta agonist combination. It's really apples and oranges and that's my point.

CO-CHAIR WEISS: Yes, and I think that the -- the parallel thing we are hearing is, is that the measure developers were using the guidelines as a way of demarcating these categories and that they depended upon the guidelines as a source of evidence summation.

And what I'm hearing from you is, is that there's some concern with how you see the guideline, the national asthma education program guidelines have summated in terms of how to use it in this measure. Is that kind of what we are hearing?

MEMBER LANG: Right, but the -- the third expert kind of report guidelines do stipulate that long acting beta agonist should not be prescribed as monotherapy. I mean that's a clear message and that medication is here, you know, if it's prescribed in

combination with inhaled corticosteroid, that's different.

But I don't see that your guideline allows us to discriminate those two events necessarily.

CO-CHAIR WEISS: Okay, so let me just be mindful, because I don't want to get into a long discussion on evidence, but you're raising the issues that I think are salient for us to be considering as we think about validity here.

Any comments on what David has said about his concerns or any other comments with regards to validity? Mark, final response because we have to move on. But please do, if you can keep it brief.

DR. ANTMAN: As succinct as I can. Once again I apologize that we recognize that we did have -- that is an error in our definition for the numerator.

Long acting inhaled beta2 agonists and short acting inhaled beta2 agonists, are

not supposed to be in that list. So that was an error. We corrected the specifications but not the language of the definition.

CO-CHAIR WEISS: Is that a moment of never mind, or? Still concerned, but not on that issue.

DR. ANTMAN: Well, it does --

CO-CHAIR WEISS: Not on that very specific issue of --

MEMBER LANG: Well, that was a major concern, is what you just said. So those agents, again, just to reiterate, just to be absolutely precise here, if a patient then receives a long acting beta agonist prescription, and that's the only prescription they receive, how is that handled?

DR. ANTMAN: The measure is not met.

MEMBER LANG: Okay, and if patients receive a short acting beta agonist and that's it, they also don't fulfil the measure.

DR. ANTMAN: Correct.

MEMBER LANG: Okay. All right. Very good. You know, this still then has some of the similar concerns regarding methylxanthines, leukotriene modifiers which are not in the same category of evidence as inhaled corticosteroid, but we've -- those issues have been put in front of the group previously this morning. Thank you for clarifying Mark.

CO-CHAIR WEISS: Great. Okay. So let's then vote on reliability, one through four.

(Pause for voting)

CO-CHAIR WEISS: Looks like we got -- 5 say high, 15 say moderate, no low and no insufficient. Next we'll go to the more debated, validity, one through four. Please vote.

(Pause for voting)

CO-CHAIR WEISS: Let's all just press our buttons again and send, just in

case. There we go. Ooh. No highs, 14 moderates and 6 lows, no insufficient, so it still passes but not very enthusiastic.

Okay. Let's go to the usability and feasibility. So, David?

MEMBER LANG: Yes, I think that the -- I can address them both together in terms of time. There are no major issues regarding usability per se. I mean I think that the -- I think the measure has been -- the reliability of the measure is -- excuse me. The measure has been tested for feasibility. Again, we are dealing with largely electronic data, pharmacy claims and a definition of persistent asthma. So I think we're good on both, and in the conference call that was reflected in the votes.

CO-CHAIR WEISS: Members of the workgroup, any other comments to David's?

(No response)

CO-CHAIR WEISS: Okay. Workgroup at large, any questions, thoughts, concerns?

Let's vote. Usability, one through four.

(Pause for voting)

CO-CHAIR WEISS: Okay. Let's see what we've got, 11 say high, 7 say moderate, 2 low and no insufficient. Next we go to feasibility, one through four again.

(Pause for voting)

CO-CHAIR WEISS: Hold on to your thing. We are going to vote for summative. That's the final piece here. Okay. Who's that 20th person? Let's all press again.

There you go. Okay, 11 high, 9 moderate, no low and no insufficient. And finally to the summative overall, yes/no.

(Pause for voting)

CO-CHAIR WEISS: Sixteen yes. It passes. Let's go on to next measure. We have trimmed about 30 seconds off the last one.

CO-CHAIR WEISS: Thanks so much. So, Denise, from the Joint Commission as our measure developer. Please.

MS. KRUSENOSKI: Good morning, I'm

Denise from the Joint Commission, I have with me Ann Watt here as well. We have three -- and on the phone we have Dr. Nimmagadda and also measure developer Elvira Ryan.

We have three pediatric, inpatient measures, 0143, 0144, 0338. These measures have been collected since 2007. They are publicly reported on Hospital Compare and on the Joint Commission's quality check website.

All of these measures are in the process of retooling for electronic collection, and they are included in the proposed rule for stage two of meaningful use.

The first measure, 0143, is stratified, ages 2 through 4 years, 5 through 12 years and 13 through 17 years of age. This first measure looks at the use -- it's a process measure looking at the use of relievers for inpatient asthma.

CO-CHAIR WEISS: That's it?

MS. KRUSENOSKI: Would you like me to go to the second one?

CO-CHAIR WEISS: Yes, why don't you do all three, if that would be okay, if -- do you feel like you can or do you want to keep them separate? What would work best for you?

MS. KRUSENOSKI: Sure, no this -- no. I will continue as well.

CO-CHAIR WEISS: That'd be great.

MS. KRUSENOSKI: The second measure, 0144, is looking at the systemic corticosteroid use of again, inpatient, asthmatic, pediatric patients. It's stratified with the age groups as well.

And the third measure is 0338, which is the home management plan of care document given to the patient or the caregiver, which is an individualized, written plan of care.

It's personalized to the child, specific to their followup care, their identification of triggers for their asthma, a rescue plan that's been identified for that

child, use of their home medications, and evidence that this document was presented to the family and then evidence that it is present on the chart. Those are the data elements for that last measure, 0338.

CO-CHAIR WEISS: Great. So we are going to look first at impact, opportunity and evidence, and Trude. I was looking for your first name. Trude, thank you.

MEMBER HAECKER: This is obviously something that has been used for many, many years and so the impact, asthma is clearly the number one diagnosis for chronic disease states in children. Can you not hear me?

CO-CHAIR WEISS: Bring the microphone real close.

MEMBER HAECKER: Sorry.

CO-CHAIR WEISS: Make it a friend.

MEMBER HAECKER: Steal it from you. So, asthma is the number one chronic disease of childhood. Rates of asthma, you know, correlate quite highly in the inner

city. In Philadelphia we have 22 percent of kids with asthma.

So the use of -- the impact of this, we felt, as a group, was quite high. So no concerns there. Rationale, there's been years of evidence of the use of relievers in inpatient settings so we also had no qualms about that as well.

Do you want to keep going? Scientific acceptability --

CO-CHAIR WEISS: Opportunity.

MEMBER HAECKER: Opportunities, I think are very limited. That's where we, I think the group was -- because we have 99 percent rates already, so we are doing very well in those children's hospitals, those of my colleagues in the room, so that it is a wonderful measure, it is useful, we report it, but again, we are doing this as part of our care routinely. So --

DR. WINKLER: I checked Hospital Compare yesterday. The national rates are 100

percent.

MEMBER HAECKER: Exactly. We are at 100 -- we've been at 100 percent since 2008 at CHOP and I'm sure you are here at D.C. and Pittsburgh as well.

CO-CHAIR WEISS: So with that in mind, to the group, the workgroup, thoughts or comments on Trude's comments?

MEMBER GLOMB: If I can elaborate, just -- she's dead on. I think this is a -- it had its place and time and the impact has been made. The impact was necessary but we have really swung far beyond it.

She's citing 100 percent rates. I'd claim 110 percent rates because we are overtreating from a specialist standpoint and even from a payer standpoint, you know, we've really gone the other direction.

I think that the literature cited for this measure now is ancient history, particularly in the evolution of pediatric asthma diagnosis and treatment, and it may be

a retirement.

CO-CHAIR WEISS: So a question I would have is we are moving towards the concept of reserve, when does that happen in our process?

MEMBER GLOMB: We should first vote it down and then ask it for reserve.

CO-CHAIR WEISS: Okay.

MEMBER HAECKER: The other piece to this is the issue of electronic health records, and so order sets are being created now for asthma in most institutions.

So this is part of every order set electronically as well so that actually keeps you at 100 percent no matter what.

CO-CHAIR WEISS: Comment. Maybe if you can slide over to another microphone and see if you can grab something that way. Folks on the phone, we have --

MS. WATT: Sorry my name is Ann Watt. I'm from the Joint Commission. And obviously, we can't argue the fact that this

measure is being met at a very high rate.

Just one thing though that I want to point out for you, is the hospital -- it's is a relatively small group of hospitals that are reporting on this measure, and we feel that it is the group that -- for whom this is a particular concern.

And what we think is, it's a self-selected group, not necessarily representative of general hospitals as a rule. We would like for this measure to continue to receive its active endorsement, just because we feel that the opportunity is bigger than the small group that is currently reporting, and assuming that it does move forward for meaningful use stage two, there will be plenty more hospitals reporting on it whose rates may not be as high now.

CO-CHAIR WEISS: Is there any evidence of that, of the non-reporting hospitals in terms of this measure, because it is viewed as pretty much a standard of care

that's been pretty well embedded in. I mean, is there any hospital that there are hospitals who don't have high rates here?

MS. WATT: Well, because we only have reporting hospitals and their rate is high. But again, they are a self-selected group.

CO-CHAIR WEISS: Brendle, I want to keep this relatively short.

MEMBER GLOMB: Yes, I don't want to belabor the point either. I have no evidence to my point than she has for hers. But I can tell you that in a big state, vast rural areas, even our most unperforming hospitals in the state of Texas are performing -- who certainly wouldn't be reporting -- are performing at 100 percent, or close to it.

CO-CHAIR WEISS: So, at least anecdotally we are hearing that, and it sounds like, with the information we have right now, in terms of evidence, that it looks like there is no performance gap to clear up although

there is a hypothetical one that we would like to see, but we don't have that information. Okay.

So, let's go through the vote of impact, high through insufficient. Let's vote.

(Pause for voting)

CO-CHAIR WEISS: Let's vote. Press your numbers again, just in case. There we go. Okay. Got it. So 13 said it was high impact, 3 moderate, 4 low, no insufficient.

Let's go now to the more discussed issue, which is the performance gap. So how many view this as a high, moderate low, and then insufficient evidence?

(Pause for voting)

CO-CHAIR WEISS: Looks like we got them all. Okay. So this is 1 high, 1 moderate, 18 low, 4 insufficient. Does that stop us here? Sorry, zero insufficient. I keep on doing that. I'm sorry. That stops us right here? Okay.

So we go straight to a question. So now, so essentially we have said no to moving this forward but we can have a conversation of reserve and there's no voting, electronic voting for this, but we can vote.

Oh, you are good. Look at that.

DR. BURSTIN: Yes, we are --

CO-CHAIR WEISS: So now we go to reserve status, and maybe since it's the first time, maybe Reva, if you can just give us the -- anything you'd like us to know because --

DR. WINKLER: Reserve status is for a stellar measure that is performing extremely highly and must meet all the criteria very, very highly -- strong direct evidence, proximal to the desired outcome, high ratings for reliability and validity, it's demonstrated in use and demonstrated improvement.

So there really can't be anything questionable or concerning about the measure. But if indeed you feel that it is of such high

import that you want to keep it on NQF's list of endorsed measures, albeit on the reserve shelf, such that it could be pulled out for later use for either, maybe new hospitals joining the party, or for double check in a couple of years to see if there's been any backsliding, those would be your rationale, as opposed to just letting the measure go.

CO-CHAIR WEISS: So, any questions to Reva on that concept?

MEMBER HAECKER: So could you clarify that again? So a yes implies that the measure well, would go into reserve?

DR. WINKLER: Well, that's what -- in order to finally put it in reserve, we would then have to go through the rest of the criteria because we have to be sure it does meet the others highly.

But yes, essentially yes means you want to consider it for a reserve status. Saying yes we know it doesn't meet the gap, performance gap sub-criteria, but we are

making this special exception to put it in reserve status. So that's what your yes vote means.

MEMBER STEMPLE: Does yes mean we anticipate or concern for decline? Because I want to make sure, because if I'm hearing everybody's thinking it's in the electronic set, so the potential for this to underperform going forward is probably pretty low, so is that part of the reserve criteria, our risk for underperformance is anticipated or --

DR. WINKLER: Yes, the primary rationale for reserve status is that concern, that going forward there could be reduced performance, and you'd want to be able to have a tool to, to measure it again. That's the primary rationale.

MEMBER JEWELL: So I think in this case it also is the question of, since we don't have data on many hospitals because they are not participating, the potential that we'd discover lack of performance would be the

other thought.

CO-CHAIR WEISS: Okay. So let's vote right now just to say that we want to consider. Yes, no, one, two. Yes being yes, let's consider it for reserve status, no being no.

(Pause for voting)

CO-CHAIR WEISS: Overwhelmingly, 18 say yes, let's consider it. So it's going to do -- consider it, now we have to go through the process of consideration? Okay.

DR. WINKLER: Yes. Now you'll go back and hit the 1c evidence vote and then the rest of the -- right. There.

CO-CHAIR WEISS: Oh, okay. So is there evidence? Yes, no, insufficient evidence that this is a good measure. Yes, no, vote now.

(Pause for voting)

MEMBER LEVY: Why would you go through this if we already decided it should be reserved? Aren't we already saying that

that's true, by saying it's in reserve?

CO-CHAIR WEISS: I guess, Mitchell, the question is did you vote, and then can we talk about this at the same time?

MEMBER LEVY: Yes I did.

CO-CHAIR WEISS: Okay, I want you to vote. So, one, two, three, please make your vote because only 19 people have voted, and let's respond to Mitchell's question.

DR. WINKLER: The -- because your vote was to consider it for reserve status, it cannot be voted on reserve status until it meets all the other criteria we haven't voted on as yet.

DR. BURSTIN: So, essentially the idea would be we wouldn't want to put it into reserve status something that you don't think is highly reliable or valid, for example. In that case it should just be removed from endorsement, which is the other choice.

CO-CHAIR WEISS: So let's go back

and make sure everyone has voted. One, two or three, let's hit your buttons again. We are so close to doing well on timing. We still may get through this one in a reasonable time.

There we go. Somehow we got the last one. Okay. Evidence is strong, 20 yes. Next.

Okay. Reliable, high, moderate, low, insufficient. Do we want to have a discussion on this?

(No response)

CO-CHAIR WEISS: No. Okay. So let's vote.

(Pause for voting)

CO-CHAIR WEISS: Everyone make sure you hit your button again please, and maybe point it to Jessica. She is wanting the attention. There we go. You see that helped, 19 high, 1 moderate, no low, for insufficient -- just kidding -- no, no insufficient evidence.

Next. Validity, this was the

validity measure, one high, two moderate, three low, four insufficient. Anyone want to discuss anything here before we start going voting?

(No response)

CO-CHAIR WEISS: Okay. Then let's vote.

(Pause for voting)

CO-CHAIR WEISS: Let's press them again everybody. It could be that one has got a low battery and it will be impossible to find it. So one more time let's all point to Jessica.

There we go. Okay, we got it, 17 high, 3 moderate, no low, no insufficient. Next. And this would be usability. One, two, three, high, moderate, low, and then four insufficient.

(Pause for voting)

And we'll come back for feasibility. Oh, so close. You'll hear it.

(Alarm sounds)

CO-CHAIR WEISS: That's the 15-minute mark. Again, let's -- got it, okay, and then we'll go to feasibility. So 13 high, 6 moderate, 1 low, no insufficient. And now to the last, usability -- feasibility, sorry. High, moderate, low, insufficient.

(Pause for voting)

CO-CHAIR WEISS: Okay. There we go. And we're at 18 high, 2 moderate and no low, no insufficient. So I think over suitability for endorsement for a reserve measure, yes or no, this is the final, final. Please vote.

(Pause for voting)

CO-CHAIR WEISS: Did I do that right? Did I do something wrong? Make sure we vote, one or two. Press yours again if you could, everybody. Okay. Almost there.

Almost there. Let's go one more time everybody. Press them down. Smile, Jessica. It's all coming to you. There you go. That big smile made a difference, see?

Nineteen yes, one no.

Okay, two minutes over, but that wasn't bad, right? We did set a benchmark for reserve status. That's true. It's going to be hard to beat that one.

Next. So 0144. Brendle.

CO-CHAIR WEISS: We already had introduction, so the question for you is on impact and gap and opportunity. I'm sorry, gap and -- thanks.

MEMBER GLOMB: All right. So just as a refresher for everybody, this is looking at, again, a pediatric inpatient drive, systemic corticosteroids during hospitalization in percentage.

Looking at impact, I believe it's recognized by the entire subcommittee that this is a -- if not a health outcome direct measure, plenty of evidence that there's high impact for this measure, large substantial impact, and that -- we're not looking at, we're not looking at evidence yet, right?

Just still impact?

Okay, just impact. All three. Okay, all right. That there's considerable evidence to the positive with regard to its favorable from the long -- standpoint and potential for benefit compared with the potential for harm, burden, and that there are a considerable number of studies, they're all relatively -- very good studies, difficult to completely account for confounding variables within most of these studies, short acting bronchodilator, administration, oxygen application, epidemiology and causation of the exacerbation et cetera, but the literature is fairly uniform in its results, in its findings.

The gap is small but there is still opportunity within that.

DR. BURSTIN: It's 98.8 percent.

DR. WINKLER: It's 100 percent also on Hospital Compare's of yesterday, the national average.

CO-CHAIR WEISS: Thoughts or comments from the workgroup?

MEMBER YEALY: So my question would be, is this a recently achieved 100 percent mark, or persistently? The last one, it really had been persistent, and where it went to reserve, that would be my question.

CO-CHAIR WEISS: What's it been, the past two, three years? What does it look like is what you're asking?

MEMBER YEALY: Last two to three compared to previous.

MS. KRUSENOSKI: Sure, in 2007, performance was at 97.1 percent, and second quarter of 2011, 99.3 percent.

CO-CHAIR WEISS: Mitchell? No. Okay. So it sounds like -- no. Dianne?

MEMBER JEWELL: Are we talking about the same number of reporting hospitals, for Hospital Compare, or roughly the same?

DR. WINKLER: Yes, exactly.

MEMBER JEWELL: Okay thank you.

CO-CHAIR WEISS: Thoughts, comments, questions?

(No response)

CO-CHAIR WEISS: Okay, let's go to vote. So we are voting on impact.

(Pause for voting)

CO-CHAIR WEISS: And we're voting again on impact. Make sure you press the number and then -- there you go, 18 say high impact, 2 say moderate, no lows, and no insufficients.

Next. Let's vote on the gap, which is somewhere in the high, high 90s.

(Pause for voting)

CO-CHAIR WEISS: Let's vote again on the gap. And again, until we get the answer we want. There we go. Okay, so 1 high, 4 moderate, 15 low and no insufficient, which would mean it would not pass. So we would go now to the question of reserve. This feels so sad to have success like this, doesn't it?

So reserve, should we consider it

for reserve, yes, no. Let's vote on that, unless anyone has another question about that. Okay, let's vote on it.

(Pause for voting)

CO-CHAIR WEISS: Okay, got 20, and it's twenty that says yes. Okay. So let's go through the reserve process. Let's continue forward. So --

DR. BURSTIN: Can we have just one question since it's the exact same methodology, reliability, validity, usability, feasibility, I wonder if we could ask the committee if they want to --

CO-CHAIR WEISS: So let's ask the committee.

DR. BURSTIN: -- the same way and then just go straight to approve reserve status. There you go.

CO-CHAIR WEISS: So you feel like with that -- let me make sure that from the measure developer, is there anything you'd like to comment on before we go to vote, just

so that -- because we are going to --

MS. WATT: Yes, we agree, that this is exactly the same methodology, same hospitals collecting the data and so forth.

CO-CHAIR WEISS: So the real opportunity might be if we see some gap with hospitals outside of that network, that this would be able to be pulled off the shelf. Okay. Good.

So with that in mind, straight to the last vote.

DR. BURSTIN: This is would it be suitable for reserve status endorsement.

CO-CHAIR WEISS: Okay, so that's it. One or two. Yes or no. Reserve yes, or not.

(Pause for voting)

CO-CHAIR WEISS: Twenty say yes. Okay. We broke our benchmark. Okay, picked up a little bit of time. Good let's continue on to number three, Measure 0338. That goes back to Trude.

We got, we gained some. Bunches of minutes.

MEMBER HAECKER: I'll use them up now. So this is home management plan of care, which we -- just to remind everyone is based on admissions to the hospital for children under the age of 18 as a primary diagnosis of 493.

There are five criteria, as you see on your handout there, in addition to their measure being -- require that you have documentation on the chart, and also documentation that the care plan was given to the family.

So that's really seven measures. There's also an all or none measure, so all those criteria need to be met in order for it to be acceptable to the Joint Commission.

It has also been set up that that is a benchmark for surveying so when they walk into a pediatric hospital, you know, you need to meet a benchmark of 80 percent in order for

them to continue their survey.

So it is a quite stringent requirement, which has driven, I think the numbers up. So you can see there the numerator and the denominator, and I think one of the conversations that the workgroup had was that this becomes more of a work flow issue, getting residents and house staff and others to work on this. While we all applaud education tremendously and it's highly important to be able to educate families at the point of care, and to make sure that they are leaving the hospital with a real understanding of what to do, and that management plan of care has clearly what to do in a flare, what to do for daily maintenance and what to do when you are in trouble, and also who your provider is. So it really fits into the PCMH, medical home issues as well.

I think some of the recent data has suggested that it may not be as helpful in some populations. So I'm going to look to my

colleagues down this end of the table to help me out with this as well.

So I think we can go, scroll down, and talk a little bit about impact, I guess. So clearly again, asthma is the most common diagnosis of childhood, and I think we all felt very strongly that the rationale and the impact of this was quite high.

I think the evidence, there is evidence out there, of the importance of education and using an asthma care plan, it is in the NLBH guidelines as well, and I think the -- there's some concern, and I'll ask my co-panel members to come into this, about the quality of the care plan. So this is the rub. There's no standardization of what a care plan looks like. Necessarily it has to have those elements, but how the language is constructed, there's not an opportunity always to have health literacy issues in there as well as there's not always an opportunity to have it in multiple languages. So I might say that

that was a point of discussion from our group.

Keep going? Okay.

CO-CHAIR WEISS: Gap.

MEMBER HAECKER: So as you can see this is part of the 3-CAC measure and we do have a gap there, so it's 79 percent is what we're demonstrating across, even in our own hospital, we are at 85 percent, so we have not reached 100 percent.

Again, some of the technical challenges of creating that in an electronic health record, and catching the patients before they leave, I think.

And then I would say that quite honestly the rub is documentation. The nurses are teaching. We are all feeling like we are teaching. But getting that documentation and an actual copy in the chart, if those criteria are not met, this metric is not met.

CO-CHAIR WEISS: Workgroup? Comments beyond the --

MEMBER GLOMB: Two brief ones. I

was a little hesitant at first reading this with the degree of specifications that had to be met. But they are all certainly appropriate and welcome within the guidelines.

Second comment, I think that, reading the history on this measure, I think there were some appropriate tweaks made along the way, particularly with regard to patients who were from out of town, patients who are being discharged on a weekend, all of these sorts of things. There were allowances made as long as the plan -- the discharge plan included the ways of getting to the ultimate goal.

MEMBER HAECKER: The other thing that was changed was the followup plan. Initially we were required to give a date, time and appointment for the followup, and that was the issue of the weekend coverage, how do you get an appointment for someone when the office is closed on Sunday afternoon at 4?

So that caveat was met because you

can now just talk about who the primary provider is, though the primary office with a phone number.

CO-CHAIR WEISS: David.

MEMBER LANG: The other view, as was mentioned, but just to embellish that a little bit more, is the possibility of the variable quality of education. I mean, education is a good thing, I mean this is Mom and apple pie kind of stuff, but although there are -- you know, some of the data supporting the utility of asthma action plans is not as strong as with, say, you know, inhaled corticosteroids as long as that was mentioned earlier, and it was relevant to our previous discussions.

But I think in terms of evidence and validity, you know, the issue is the variable nature in which this information may be relayed, and the documentation of that, and also, if I can extend -- I guess this goes all the way to feasibility -- the issue of

retrieval of those data from either an electronic or even paper record.

But even electronic, it's not -- it's not the same as retrieving, say, a prescription dispensed for drug X.

CO-CHAIR WEISS: I'll take my co-chair hat off for a moment and just be as a member of the workgroup. I think we did also discuss the leverage piece here, how much of a lever was this unto by itself, for really demonstrated that there -- just having documentation of this kind of a plan being given, showing that it actually improved any outcomes, and is it more than the plan, is it ensuring the transfer into the care process and the follow-through and all those other pieces as a comprehensive -- did I capture that right?

MEMBER HAECKER: That's absolutely correct. I think the evidence on followup and keeping that appointment back in the medical home is a much better predictor of what this

would be about, rather than just giving a piece of paper.

CO-CHAIR WEISS: Dianne, and then Chuck.

MEMBER JEWELL: So, my first read on the specification was where's the item that says evidence of understanding, which you know, of course is related to that.

However, the fact that there are five or six items itemized out specifically, that at least there has to be a category discussion, whatever quality it is, and if we're not able to keep track of those five things and the guidelines are clear about them, my head space is back to where at least we are getting everybody to talk about the same things consistently and maybe that's where we still need to be.

CO-CHAIR WEISS: Chuck.

MEMBER STEMPLE: Well, and maybe it's -- so the age cutoff at two was my number one question, why two is used as a cutoff and

not down to lower.

And then to your point, you know, return to admission status, so a readmission rate based on whether this had an impact.

So for me, less did they follow up with the office visit, but are we impacting the readmissions in asthma patients who got a discharge plan versus those who did not.

So I would further your evidence that this would provide a differential outcome, but I would lean more toward a return to ER visit or readmission rather than just a followup office visit as a downstream medical outcome to really validate that this is having an impact.

CO-CHAIR WEISS: To the group as a whole, you heard it from the workgroup there. Let's go to David and then to Norm.

MEMBER STOCKWELL: Did the workgroup, did you guys discuss the article that showed that there was no association

between the completion of the asthma action plan and ED visits and readmissions?

I mean that's been our big -- this is a process measure obviously, but does the process measure represent anything to do with the outcome, and that's what the worry is, I think, is that it may not have anything to do with the outcome, and boy do we spend a lot of time, I'm sure all of us do, on collecting the information for this.

And so if it's not -- if it's not representative of what we're trying to achieve, is it the right metric?

MEMBER HAECKER: It's kind of hard to say that I agree with you, we did bring up that article that came out this last fall and I do think it's kind of hard to fly in the face of mom and apple pie, as was said earlier, because education is never a bad thing.

But the efficacy of what we're doing and the processes involved, as you said,

many systems are hiring one full-time person to manage this process.

We have 3,000 asthmatics admitted every year. That's maybe a good thing or a bad thing, but clearly the work that's involved around that is quite cumbersome.

CO-CHAIR WEISS: Norm.

MEMBER EDELMAN: Yes, I was just going to summarize what I thought I heard, just to make sure I am right. So it sounds like the reliability is low and the validity is wholly unproven. Am I correct?

CO-CHAIR WEISS: Over here at least the validity is measured by at least one study, so that there wasn't proven, we don't know wholly unproven yet.

But that was generally the term that we were hearing, the general direction. Brendle.

MEMBER GLOMB: I was just going to comment, I think it does go to due diligence, that this is a measure that goes to due

diligence. It's the you can lead a horse to water concept, and I think that this is a way -- used in the appropriate way by a hospital facility, they can complete their due diligence to the patient, "complete" being in quotation marks, with this measure.

But understand --

CO-CHAIR WEISS: Mitchell.

MEMBER LEVY: And I just want to make sure I understand this. So this is a joint measure that's already been collected.

MEMBER HAECKER: Yes. Yes. For several years now. So we are collecting this data and we started out at very low points and we have all been working our way up into the 80s, not quite at the 90s, has to do with the turnover rate as well. I'm just giving you all the permutations on this.

Asthmatic patients stay an average length of stay is about a day and a half to two days, at the most. So you are doing -- and you're sending them home evenings,

mornings, all kinds of times.

CO-CHAIR WEISS: Yes, question.

DR. BURSTIN: So since it is, you know, one of the questions on evidence, this specifically for process measures, is there a link to outcomes, the point that you've raised? I'd be curious -- and do you guys have any other evidence to cite of that process outcome link?

MS. WATT: We don't have specific evidence to cite to make that link, but what we do have, and I'm going to ask Dr. Nimmagadda if he is on the line and if he could perhaps address this too.

But you know, one point I would like to make about that one study, it was one study that has been done, and it looked, some of the specific limitations noted in the study was that it didn't look at severity and those kinds of issues.

And so I guess I would ask that you consider that when you are considering. I

don't know. Dr. Nimmagadda, did you have any comment? Are you there?

CO-CHAIR WEISS: Dr. Nimmagadda are you there? You mean -- he may not be there. Chuck and then Mitchell and --

MEMBER STEMPLE: So, I'm sorry, I just want to make sure, we have one potentially negative study but no -- so there's only literally on study on this as an outcome? So one negative with maybe some limitations but no positive that this has caused an improvement in downstream outcome?

DR. NIMMAGADDA: Hello.

CO-CHAIR WEISS: That might be Dr. Nimmagadda?

DR. NIMMAGADDA: Yes, yes, I'm here.

CO-CHAIR WEISS: Welcome. Ann, do you want to pose the question to him again?

MS. WATT: Hi Dr. Nimmagadda. This is Ann. There was some question of whether or not we have specific evidence for

improvement in outcomes based on the home management plan of care. The discussion was begun by a discussion of the one study that was published last fall that although there were noted to be limitations to the study, indicated that there was not a link to outcomes.

DR. NIMMAGADDA: Yes, I read that study, and you know, I have a lot of questions about what that -- over that publication. One, it doesn't really show the effectiveness of what was taken into that measure.

On other words, if they just checked the boxes but they didn't really go through the processes of identifying all the components of that measure, then yes, then it's hard to prove the outcomes of that.

Also, within a pediatric population, we've seen numerous outcome studies looking at asthma action plans, and peak flow plans.

There's a difference between peak

flow plans, compliance measures with that, versus non-peak flow plans and all. But the, you know, this measure here really looked at different components in trying to do the outpatient visit, they can look at the oral corticosteroid, they can look at, you know, the controllers versus relievers, educating the patient on those components, and also taking a look at the environmental triggers here.

So this data that was published, I really have a lot of questions about, because there's really no real confirmation they actually did what they were supposed to do to make the -- this measure effective.

So it's hard to say that, you know, you can say one study proves against it, but I'm sure that if you take the components out individually and take a look at each individual outcome here, that this measure did improve outcomes, from the pediatric perspective.

Now, if you take a look at the adult institutions and other hospitals, maybe you know, there may be a little bit less of a compliance rate there, or there may not be as great of an outcome measure known.

But within the pediatric institutions in the study that we've seen, that the components here would definitely reduce the readmission rates and even the ER return rate.

CO-CHAIR WEISS: Mitchell.

MEMBER LEVY: So I feel like I'm getting a mixed message. I'm not sure. Because now you are saying there are -- there is a relationship between the process measure and the outcomes.

But my main question is I'm surprised, because usually the Joint Commission before it releases a measure, also has a rigorous process of looking for a relationship with outcome.

So is it, is my understanding

correct that this measure has never been linked to outcomes, for a pediatric population?

DR. NIMMAGADDA: No, it has been linked to outcomes in pediatric populations. But I think -- I thought the question was related to the one publication that was presented last fall.

CO-CHAIR WEISS: So now I'm confused as well I think. You're hearing a little bit of discussion around the committee because we are trying to understand this better.

MEMBER HAECKER: Yes, I think the use of asthma action plans, asthma care plans, home management plans again, has evidence in the outpatient setting.

And so within the context of the medical home, we actually give them out all the time and use them, and that data has been clear.

What I don't think we have yet is

the data of the -- for the patient that has been admitted with those different classifications of asthma, perhaps not the --

(Alarm sounds.)

MEMBER HAECKER: intermittent asthmatic that comes in, do we have data on that patient, and that's what that study recommended.

CO-CHAIR WEISS: Ah, good. First of all, everyone knows that that's the 15-minute mark.

MEMBER RHEW: Just a few comments here also. I have actually been looking at our database and there are multiple meta-analyses and systematic reviews on this topic, AHRQ (2001), Gibson (2002, 2003), I mean there's extensive literature out there and the consensus is that just handing the plan, or having that written document does not impact the outcomes.

But if you're talking about an overall program in which there's extensive

education delivered, and this is a part of that, then yes, there is a benefit.

So you know, the question is if you are looking specifically at this document, then the answer is no, no outcome. If you are looking at an overall approach in which this could be a component, then yes.

CO-CHAIR WEISS: I have to say that's how I understand the literature too. It's never been demonstrated or isolated as a management plan by, in an -- by itself, and all the work has been done in the ambulatory arena.

But Dr. Nimmagadda, do you know of any studies on discharge from inpatient looking at management plan, which is the measure at hand, by -- in and to by itself, even if it's done well, that would show positive impact? Because that's what I think the committee is looking to hear. Is that -- we only have the one study that's the negative right now, and do you know of any studies?

DR. NIMMAGADDA: You know, I do know some plans that are -- that have different components of the CAC-3 measure. There are -- there's a study that shows the peak flow plans are effective in identifying exacerbations that may start up early and take through discharge.

There's also studies in outcomes looking at the environmental control measures and the identification of patients who smoke, and you know, going back to home smoking in places, and also the kids, when they get discharged.

So I don't know if there's a study that looks at a very comprehensive plan such as a CAC-3 have got, different components in there. But the individual --

CO-CHAIR WEISS: I think the question --

DR. NIMMAGADDA: are definitely --

CO-CHAIR WEISS: I'm sorry to -- because we are running short on time so I'm

being a little bit directive in my questioning, with apologies.

But the question is that the literature that we understand is, is that it's the plan in the context of an educational activity that is one --

MEMBER HAECKER: In the context of a medical home.

CO-CHAIR WEISS: Of a medical home, would be the more recent context, but we don't know of any studies that just show just the use of a document in a one-time event, really has an effect on outcome. That's what I think the committee is struggling with here, at least those who know this literature, and I'm getting a lot of affirmative nods here. So are we --

DR. NIMMAGADDA: From an outpatient setting, yes, I mean, we have seen that these documents do have an impact in aftercare, reducing quality of -- or increasing quality of life, which is the

asthma morbidity.

But we try to take these outpatient processes and put them into an inpatient type of a setting. So there's numerous studies looking at the outpatient setting, but very few looking at the inpatient continuum.

So that's why we are trying to get this measure implemented here, to try to bridge that gap that we have with the inpatient/outpatient arena.

CO-CHAIR WEISS: That's very helpful. Good. Let me suggest, unless there's any other additional questions, I think we are ready for a vote, yes? Yes? Okay. Good.

So let's do the vote on impact, one through four, high, moderate, low insufficient. Let's all vote.

(Pause for voting.)

CO-CHAIR WEISS: It's been a while since we've done it so people are like, how do we do this thing again? Come on, there's 19,

there's 20. Perfect. Okay.

So six high, nine moderate, two low and three insufficient. Next go to the gap. Okay is there a gap in practice here that can be fixed?

(Pause for voting.)

CO-CHAIR WEISS: With regards to performance gap, seven high, 12 moderate, no low and one insufficient. Let's go to the one that I think has been the most discussion, which is evidence. Is there sufficient evidence related to outcomes, quantity and quality of the evidence, and that's yes, no or insufficient.

(Pause for voting.)

CO-CHAIR WEISS: Seventeen, 18 -- okay everybody let's -- oh there, we're all set. So I think we have hit a four yes, six no and 10 have insufficient, which puts this into a done.

Okay. Measure fails. Okay. Well thank you -- all -- okay, so that's it. With

a -- okay. So deep breath everybody. We want to thank our colleagues from the Joint Commission. It's been a tough morning for you all in the sense that you have succeeded with two measures beyond anyone's wildest imagination, and this measures as we have evolved, has -- creates an opportunity.

I think the opportunity that I have heard is that there's a lot of interest here and that if the measure can be looked at and thought of in the context of a more comprehensive set of -- more of a composite look at the process of discharge through transition into ambulatory care, and the success of that, I think that the committee was moving towards that is what they are looking for, I think it's what the commission might be looking for too, and it's -- and so I think that the concept that was okay a few years ago, we are hungry for more of a comprehensive type of measure to get there.

Is that -- does that reflect where

-- what I was hearing? I'm getting enough affirmative nods. Would anyone like to say it was something different than that? Brendle?

MEMBER GLOMB: Not anything different, but it was, if I can speak for the pediatricians, it was with great angst that I pressed the insufficient evidence button.

This is something we'd all like to see. I very much see this as a measure but we've got to back it up if we are going to make it scientific.

CO-CHAIR WEISS: So the commission, I think what we could say is -- well if you could think about this some more and bring forth something that was more than just the measurement plan measure, that would probably be well received. So if that's a --

MEMBER HAECKER: And I think a lot of the colleagues in the room would be willing to help with that process as well.

CO-CHAIR WEISS: Good, so it's a very positive no.

MEMBER HAECKER: Absolutely. Absolutely.

CO-CHAIR WEISS: If that can be said that way.

MEMBER HAECKER: As positive as a no can be.

CO-CHAIR WEISS: It's a constructive critique, I guess, was --

MS. KRUSENOSKI: We get it. Thank you.

CO-CHAIR WEISS: Thanks so much. Okay, so now we go to the SAC, Sub-optimal control and ACT, absence of controller therapy. This comes from the PQA, and --

CO-CHAIR WEISS: We have a developer on the line, so why don't we start with the developer. Who have we got from the developer?

DR. WINKLER: Do we have somebody from PQA on the line? Great.

DR. NAU: Yes. Hi, this is David

Nau from PQA. Can all of you hear me?

CO-CHAIR WEISS: Great. Welcome.

DR. NAU: Would you like me to give a quick rundown on the measure?

CO-CHAIR WEISS: If you could, one or two minutes' synopsis for the group, in any which way you'd like to, to support your measure.

DR. NAU: Sure. I've got a little bit of a hard time hearing what you're saying, but --

CO-CHAIR WEISS: Oh, so let me try -- is this a little bit better?

Not much. So David, if you could give us a one or two minute overview from your perspective, to help us understand the measure as best we can.

DR. NAU: Certainly. I'll give a quick synopsis. So this measure was developed several years ago, and was originally developed as a collaboration between PQA and NCQA, and tested with some different health

plans and prescription drug plans.

So this was built for a data environment in which only drug utilization data would be available and so it could be used for quality improvement and public reporting for prescription drug plans and perhaps pharmacies.

And the goal is, you know, as a two-part measure, first up really is to identify patients that we reasonably believe have uncontrolled or partly controlled persistent asthma by identifying those who have received at least, or more than three canisters of short acting beta agonists over a three month period.

And so the first step is to identify that group, under the premise that when a patient is consistently using more than one canister per month, a short acting beta agonist, they are most likely in need of some controller therapy.

And so the first goal was to

identify that rate for the patients using excessive amounts of short acting beta agonist, and then drill down into that population to identify what proportion are receiving any controller medication.

And so this is something that's been picked up by a few other prescription plans. URAC has just chosen to add it into their accreditation programs for PBMs and pharmacies.

So it's just starting to get used and we're drawing more evidence you know, as this is used more. But I think that's a quick synopsis, and happy to hear your thoughts.

CO-CHAIR WEISS: So any quick questions to our developer before we ask Rubin to take us on our journey? David.

MEMBER LANG: Yes thank you for that summary. I'm curious, in your denominator, you, in terms of your exclusions, you exclude patients who fill prescriptions for COPD medications and for pulmozyme, which

I understand, in terms of diagnostic overlap with asthma.

But you also are excluding patients who have filled one or more nasal steroid medications, and I'm wondering what the rationale is for that, as you know, allergy, I'm an allergy physician, and many of the patients I see with asthma, have concomitant rhinitis or allergic rhinitis, and you're -- it seems that you would be excluding any of those patients, if I'm understanding this correctly.

DR. NAU: Yes, I believe that the original reason for including that was just to ensure that the -- it's a fairly homogenous group of patients in the denominator, recognizing that you know, excluding those who are using nasal steroids could also be you know, asthmatics who we should be paying attention to.

But I think the goal was to try and decrease any false positives of

potentially putting patients into the denominator who may have rhinitis but not persistent asthma.

So I think, you know, that could be a debatable point. But I think that was the intent of adding that exclusion. It's just to try and make a more homogenous denominator population.

CO-CHAIR WEISS: Okay, so we Rubin, let's take us on our journey.

MEMBER COHEN: So I think -- we've debated a lot of the same material before the -- clearly the impact that asthma has on the community, the use of short acting beta agonists, the need for controller medications.

So I think in terms of the impact, we all agree that it's rated very high. We all agreed there was a performance gap and we all agreed that the evidence was adequate.

Questionable minor points, in my opinion, but there was no direct evidence cited from the literature, this is all based

on the NHLBI guidelines, which of course uses its own literature.

And also we had some questions during the phone call about just the 90 days, because most of the evidence has to do with chronic lack of controller therapy.

But otherwise, I think for the first, for part one, we all agreed this was high to moderate.

CO-CHAIR WEISS: Okay, and so from the rest of the workgroup, any thoughts or comments on Rubin's --

(No response.)

CO-CHAIR WEISS: Okay. Then to the committee as a whole, questions or thoughts you'd like to ask? Questions, issues?

(No response.)

CO-CHAIR WEISS: If not, let's vote. Okay. So importance in terms of impact, one through four. Please vote.

(Pause for voting.)

CO-CHAIR WEISS: Almost there. Okay everyone, press again. Pressing on. Here we go. There we go. So we get 17 high, three moderate, no low, no insufficient.

Next would go to the gap. Yes, I'm looking for the question. So do you remember what the data was in terms of what the performance gap was?

MEMBER COHEN: I was actually looking, I couldn't --

CO-CHAIR WEISS: For our developer on the line, what is the performance of these two measures right now, and can you describe the population that have been tested?

DR. NAU: Yes, if I heard you correctly, you are asking about the current gap in performance, and what the perhaps current performance rates are on the measure. Did I hear you correctly?

CO-CHAIR WEISS: Correct.

DR. NAU: Yes. I honestly am in an airport and don't have those numbers right

in front of me, but I do know that there is a clear gap in performance and room for improvement. I would be trying to remember off the top of my head what the specific numbers were.

We've tested it with several PDMs and some health plans, and identified that there's a fairly significant number of patients who are using greater than one short acting beta agonist inhaler a month who were then not on inhaled corticosteroids.

But I don't have those numbers in front of me at the moment.

CO-CHAIR WEISS: Well, looks like we don't have any submitted for this. We don't know that?

MEMBER COHEN: I don't believe it's in the original thing that you had sent me. It's not there. That's for sure.

CO-CHAIR WEISS: Okay, so, well, then we have high, moderate, low and insufficient evidence. So let's vote.

(Pause for voting.)

CO-CHAIR WEISS: So three high, one moderate, four low, and 12 insufficient evidence. So we are stuck here. Okay. So next to our colleague at -- I believe it's David, it seems that we need to get some information on performance of the measure before the committee will feel more comfortable with going forward with this measure, and so right now, we have got to say no, by the process that we've got in place.

MEMBER COHEN: If I may comment, I mean I think it's something that we all believe is probably not being done correctly in the community, but the way the rating system is, we have to stop here, yes, because you don't have a number.

CO-CHAIR WEISS: Great. Okay. Then let's go to Measure 0620 and this comes from ActiveHealth is the measure developer. Do we have someone --

DR. VIR: Hi, yes, this is Bani

Vir from ActiveHealth. Can you all hear me?

CO-CHAIR WEISS: Welcome Bani, yes we can hear you just fine. What we are doing here in case you just --

DR. VIR: Hi, I also have with me Dr. Ajay Sharma and Rajesh Makol, one of our --

CO-CHAIR WEISS: Excellent what I would -- what we are doing here is asking at the beginning of the presentation of your measure, just if you'd like to say one or two minutes' worth of introduction to your measure.

DR. VIR: Sure, I can go over -- do you want me to give you a brief description of the measure? It's a little difficult -- it's been a little difficult to hear you guys.

CO-CHAIR WEISS: I'm sorry about that, but hopefully we are hearing you well. We'll give you some solace.

DR. VIR: Okay. So I'm understanding you want me to describe the

measure.

CO-CHAIR WEISS: Yes please.

DR. VIR: Okay. So this measure is looking for the percentage of patients with asthma who have a refill, at least one refill for a short acting beta agonist in the past year.

The spirit of this measure really is to ensure that patients have access to at least one rescue inhaler. We, I just want to clarify from the get-go, in case it comes up, the heart of this measure doesn't lie in trying to delve into optimal control from a long term care management with controller meds et cetera.

It's really just looking to see that they have as a practice at least one rescue inhaler in case of an emergency.

CO-CHAIR WEISS: Excellent. Thank you. From the committee, any general questions you'd like to ask of the developer before we get started? Otherwise we'll ask

Rubin to start.

(No response.)

CO-CHAIR WEISS: No. Okay. So let's go. Rubin.

MEMBER COHEN: Yes, I think I just have to say one thing, because we had a lot of problems with this over the phone call, so just to repeat what the developer said, this has nothing to do with asthma control. It's not about inhaled corticosteroids.

It's really, do people who have asthma have access to a short acting beta agonist, because if that issue is not understood by the committee, this is going to fail on the first vote.

So, based on that, there were some -- okay, so let's go step by step. We believe that the impact was high. The performance gap was scored by the developer as being 42 percent.

We had some issues with that number, because the age group here, I believe,

was two to five and there was as a question on how you would define asthma for those less than five years of age, and also, what would you do with people who have intermittent asthma.

Their asthma my be under control, they're not having any problems, s they would not get a prescription, and that doesn't mean that they don't have access to case, just simply their asthma is well controlled.

The other issue was, I believe, one of the brands, ProAir I believe it was, has a shelf life of about two years, so a person may get that and may keep it for two years and not need to refill it but they still have access to it.

Those were questions we raised with the performance gap being 42 percent because 42 percent sounded quite impressive actually, but it was real when you take those other things into account.

DR. VIR: Can I can address that

concern?

MEMBER JEWELL: Before you do what I'd like to do is just make sure that the committee, the workgroup -- let me just take a moment here. So in terms of impact, gap and evidence, any more comments Rubin?

MEMBER COHEN: No.

CO-CHAIR WEISS: Okay, and so what we're hearing principally in the workgroup was the

concern, seeing the gap of 42 percent raised the question of whether or not they are capturing well all the use of medicine such as people who have medicines that are long shelf life and may use it that way, and then it comes to mind as I'm thinking about that, maybe samples would also be there as well.

MEMBER COHEN: That's true.

CO-CHAIR WEISS: And then the other is that the denominator has cast it so wide that there may be people who don't need the medicine. They may actually not have

asthma, they just were having a single diagnosis or something. There might be something with the denominator or the numerator here, or the practice in the field is way off at 40 percent.

And that was the workgroup. Is that what the workgroup recalls, not recall? I'm getting some nods affirmatively. Yes.

MEMBER GLOMB: One clarification with that. I think you are sampling point is well taken, particularly if it's not a persistent asthmatic, you know, if we are seeing that patient in the office and it's an intermittent problem, we're going to sample -- they may get a prescription, so my question was, does this count prescriptions written or claims, because that patient may never again in the course of the year, need to use anything more than the sample that I've also given them on the way out the door and that might not be, you know, it might not be filled.

CO-CHAIR WEISS: And part of that is just because of that very broad definition of asthma. It's basically --

MEMBER GLOMB: Right, overly broad in this -- on this measure.

CO-CHAIR WEISS: Okay. So to our measure developer, you've heard the workgroup's thinking on this. Did you want to respond with any more information?

I'm not sure -- do we still have you on the phone? I'm not sure if we -- oh. We changed.

MS. BOSSLEY: Yes, you've got the other HB. Hi, I'm Heidi Bossley, I'm vice president, performance measures. Bani are you still on?

(No response.)

MS. BOSSLEY: I don't know, operator, can you take -- see?

CO-CHAIR WEISS: While that's going on, let me ask to the group as a whole now, because I asked Rubin, we talked a little

bit about what the workgroup thought. So what are your thoughts on this issue of impact, gap and evidence? Donald.

MEMBER YEALY: It looks to me like it's based on prescriptions filled. It says refill here. And I would have the exact concerns, the absence of filling the prescription doesn't mean the absence of good care. It can be the absence of need also.

CO-CHAIR WEISS: Any other thoughts from the committee on any of these three issues?

(No response.)

CO-CHAIR WEISS: Well let's go and vote then. 1a -- yes, do we know anything more about the measure developers?

DR. WINKLER: No, I don't think so.

CO-CHAIR WEISS: Okay.

DR. WINKLER: Bani? Just let me check one more time. Are any of the developers from ActiveHealth on the line?

(No response.)

DR. WINKLER: We could hear them --

CO-CHAIR WEISS: Okay, 1a, This has to do with impact, so our perception of impact of this measure, as a priority. Please vote one through four, one through three high, moderate, low, four being insufficient.

(Pause for voting.)

CO-CHAIR WEISS: There, we got a full complement here. Nine say high, nine say moderate, one say low and one say insufficient.

Next would be the gap. So let's vote on the gap, high, moderate, low, let's have a -- yes, we have got to get to -- before Jessica gives us that.

Okay, there's a performance gap that suggests that there's a need for improvement here, high, moderate, low, insufficient evidence.

(Pause for voting.)

CO-CHAIR WEISS: Almost there. Or is that -- maybe not. Let's make sure

everyone is voting. It has to do with performance gap, high, moderate, low and insufficient information. If you voted please vote again, and again.

Coming from Chicago, this feels so nice. If you don't want to vote you can give me your votes and I'll vote for you. I think, is that 19? Okay.

Everyone, let's point right to Jessica and so just like -- oh there we go, perfect. Okay. Three high, eight moderate, two low and seven insufficient.

It just barely passes on that. Yes. Okay, let's go and talk about the evidence. Is there evidence that's associated with the health outcome?

Yes. Yes. Yes, no, insufficient.

(Pause for voting.)

CO-CHAIR WEISS: Everyone vote again please. I'm moving to the Chicago suburbs. Oops, okay, so six says yes, one is no and 13 insufficient which means we stop

here. And the feedback to the measure developer is this concern of the denominator being excessively wide, the numerator having other mechanisms for people to either need or not need medicines, or maybe have or not have medicines, including sample and long shelf life.

Okay. Very good. That means we are up to the last measure before lunch, and it's the measure that I had the pleasure of reviewing.

So, do we have our colleagues from Minnesota Community Measurement on board?

DR. WINKLER: Do we have the operator? We don't.

MS. BOSSLEY: I've emailed the developer to find out if we just lost them or what happened, so we'll come back to that.

CO-CHAIR WEISS: Okay. Maybe since we have just one measure, do we want to just go to lunch 10 minutes --

Okay, and then we have the comment

period as well. Well, let me -- for at least -- to describe the measure, let me do it, and then what we can do is hopefully we'll have the measure developers on to ask any detailed questions on, and see if we can go that way. It's kind of like going to the very beginning of the meeting when we raised hands when we couldn't have the electronics.

So, this is a measure of optimal asthma care. It's unique among the measures that we've looked at in the asthma group because it's a composite measure, all or none, yes, no.

In order to be all yes, one has to be affirmative on all -- on three elements, and those three elements include -- I want to make sure I get this right here. What's the best, let me just find the -- thank you.

So in order for that yes/no to be a feature of the elements, one is it's well controlled by the use of one of four asthma control survey measures that are patient

surveys, and these surveys can be scored as in control or not, or well controlled.

The second part of the score is whether they have a risk of exacerbation as measured by use of emergency department of hospitalizations being greater than one.

And the third is that they have had some evidence of an asthma education and self-management with a written asthma action plan that was created and reviewed during the measurement period.

So this requires them going to chart, it requires them doing a patient survey, and also getting some level of information that can either be automated in terms of emergency department use, or collected by chart as well.

The impact was the issue, was that there was a sense of high degree of asthma prevalence, hospitalization and the need for -- - emergency department use -- and the need for measures that will comprehensively look at

asthma care.

The workgroup thought it was very clear that there was impact for the need for such a measure. In terms of performance gap, what we have from the performance measurement developer, was a very broad testing of the measure as a composite, which showed a large opportunity for improvement, and I don't have the number in front of me, but I think it was in the -- oh here it is.

Statewide adult average was 15.7 percent who actually achieved good control based upon this measure, and that was in the adult.

In the pediatric it was 24 percent. And in fact that number was raised and I was the reviewer there, and I raised to the group a concern that either the measure characteristics are concerning, or the care of asthma in Minnesota is very bad, because this is a pretty big swath of Minnesota, primary care docs.

And so I raised the question, is maybe that there's some over-specification of the measure or some mis-classification, something going on in the measure, not in the are process, that that many primary care docs in that many clinics, which is a pretty wide swath of a sample, really a hefty and good sample for this.

And actually that relates to some of the comments they gave back to us which we can reflect in a few minutes.

So the performance gap, to me, was unclear because of this low number. It looked just too big. The third part, which was the scientific evidence, was of concern in the following ways, and again I want to say that as a reviewer, I was very supportive of this idea because this was the way measures have to go in my mind, which is this composite, not looking at certain processes of care, but actually looking at multiple processes and eventually to outcomes.

And as one looked at it, and delved into it, you start to see the warts and blemishes that are needed to be looked at, and the principle one around the asthma control survey was that this asthma control survey is a survey that was developed and authenticated by performance testing in asthma clinics of allergists and -- which meant that these were areas where you have individuals who probably had a higher degree of severity of asthma and also a relationship with their asthma that probably made them good candidates for testing of survey and repeated testing of surveys.

In that environment it's a very good, reliable survey. So in one of those three elements, the survey of asthma control seemed to be good.

The difficulty is when you take it to a broad population, with a lot of very mild asthma in that population and in individuals whose diagnosis of asthma may even be relatively --

(Alarm sounds)

CO-CHAIR WEISS: modest at best. That was 15 minutes believe it or not. I better cut down my talk. Oh actually no, that was 15 minutes from the last one.

Oh well, okay good, so we haven't started yet. We are all ready for my long talk. No, just kidding. So that this asthma control instrument may not be validated in the population under study of a broad, pretty big catchment diagnosis.

So that was the concern and also getting that information back in at a high rate would be problematic when you start scaling this up.

And it turns out, on the feedback they gave us some information on that and I'll get to that in a second.

The emergency department and hospital use didn't seem to be much of a problem in terms of evidence. That seems to be pretty strong.

And then this other, third issue, the asthma management plan, seemed to be strong in terms of the literature. However it was vague to me in terms of how consistent that -- of what this means to have a management plan and education around it, because that can have a huge degree of variability, at least as they are defining it.

So I was uncertain in my mind as a reviewer as to whether these elements of the asthma control survey, which also had the additional problem of an age gap and who could be asked it, versus the parent asking it.

And then this other one about the what the management plan should be and the education plan leaving me with insufficient sense of evidence there.

So I'll stop there in terms of my interpretation. Let me ask the workgroup if I've given a reflection of what we talked about and what you all should like to say about it.

Brendle.

MEMBER GLOMB: Just a comment, I think if the -- I doubt that the asthma care in Minnesota is much different than anyplace else, better or worse.

CO-CHAIR WEISS: So it's all bad across the country, is what you're saying?

MEMBER GLOMB: Pardon me?

CO-CHAIR WEISS: So it's bad across the country? Fifteen percent?

MEMBER GLOMB: Suboptimal.

CO-CHAIR WEISS: Suboptimal.

MEMBER GLOMB: I think the elements that probably provided the biggest, because this was an all or nothing, were probably the asthma control test in a primary care setting.

I think that that is a bridge too far for some to make, given the duration of time for scheduling the patients, and then not justifying that philosophy. I'm just saying I think that's why.

And then the other may probably have to do with having all of those elements within the action plan. I think it is unusual to find an action plan from a primary care setting, probably even from some specialists, that includes both triggers and medication effects in it.

I think the other two, absolutely, but I think that would be unusual, and most primary care folks are using a pre-packaged, electronic medical record, asthma action plan -- the ones that I've seen have never included those elements -- and if not, they are downloading one from source or they are using the school's asthma action plan, and those never include those additional elements.

So I think that this was -- this is ambitious, I think was your word, David, and I think that's probably where the gap probably stems from.

CO-CHAIR WEISS: Peter and then Chuck.

MEMBER ALMENOFF: When we looked at diabetes in our system, we actually do very well, but when you try to create an all or none model, and group five of them together, we do miserably.

So you are describing the 10 percent. I wouldn't -- that doesn't really surprise me, because you know, diabetes care, we are in the 90s, but when you group five together and it's all or none, we are in the 20s or 30s in our system. So this really isn't a surprise. It's also very difficult.

CO-CHAIR WEISS: So it's that we have to be mindful in a composite you get combined probability, you know, point -- of your -- if you have 0.1, which would be 90 percent times 0.1 times 0.1 you start getting --

MEMBER ALMENOFF: And then the other issue there are, you know, there are four components some might be more -- three -- some might be more weighted than others, and

if it's an all or none and you do the one that's really important but you -- you don't do the one that probably isn't as important, you wind up failing the measure and it's sort of a -- it's a disincentive.

So just, it doesn't surprise me that you know, the composite scores are so low.

DR. WINKLER: I need to break in just to say to anybody listening on the phone, we realize we are having technical problems. We can't hear you but we think you can hear us.

So over lunch we are going to try and fix all that and give the developers for the last two measures an opportunity where they haven't been. So just to pass that message along.

CO-CHAIR WEISS: Okay. Great. Chuck.

MEMBER STEMPLE: Thank you. I thought in managed care, our risk for

exacerbation is not determined by a previous ER visit, it's non-compliance with their meds, and those people who actually had an ER visit hospitalization, returned to the norm, and they are less likely -- so you said there's a lot of data to support that so I don't know the validity of that data, but at least in my world, we would not consider an ER visit or admission a more specific risk for an exacerbation as compared to someone who is totally non-compliant with their medication.

So I don't know the data there.

CO-CHAIR WEISS: So let me help you a little bit with the data, as I understand it, which is the highest predictor for a future emergency room visit or hospitalization is a prior hospitalization or emergency visit, and that's not strictly to asthma. That's actually pretty much a utilization thing.

And however, the corollary, which is -- to that, and that is, is there a -- is

that true for most of the population? No, since most of the population will not come from that population.

So you have to be mindful that that by itself is not an all-encompassing predictor. It's a predictor for a sub-population of higher utilizers.

So let's go to David and then --

MEMBER LANG: Yes, I was just going to say briefly, because you touched on the point I was going to make, and previously we talked about components A and C, but not B, and I think that's another factor in terms of the combined probability of somebody kicking out as not being well controlled.

I think this is an issue of using a guideline definition according to the risk domain of someone not being well controlled. That is they have had more than one exacerbation as reflected in emergency department utilization, hospitalization.

But they you look at the

predictors, particularly as was mentioned by my colleague up the table, who's name is turned the other way so I can't see it, as we all are.

But the issue is that the pool -- among the pool of patients who will be in the emergency department hospital, as we all know, for the next 12 months, many of them were not in the emergency department of the hospital in the previous 12 months, and those are -- those patients may tend to be more well-behaved.

So I think this is a, this is an issue of using a guidelines definition versus epidemiologic studies that look at risk factors.

So just as long as we are passing along feedback to the measure developers, that's something they might want to keep in consideration.

CO-CHAIR WEISS: Peter and then --

MEMBER ALMENOFF: One other point I forgot to make is it's also going to be a

significant public relation issue if we are going to -- if we are going to say that we have only 20 or 30 percent compliance with a composite measure, which might not really be reflective of what we are doing and then the public will think we are doing a pretty bad job when in fact maybe we are not.

So you know I'm just a little concerned about all or none phenomenons and especially when they are not even weighted, maybe giving the wrong message.

DR. VIR: Bani Vir from ActiveHealth. We were just reconnected. Can you all hear me?

CO-CHAIR WEISS: Yes. Okay so yes we actually can --

DR. VIR: That's an ordeal we've been through.

CO-CHAIR WEISS: Yes, we apologize and we are welcoming you back. Also we would like to know, do we have the folks from the -- from Minneapolis? No, sorry, from Minnesota?

And maybe from Minneapolis.

MS. PITZEN: This is Collette from Minnesota Community Measurement, and we are now back on the line, just this minute.

CO-CHAIR WEISS: Okay that sounds great. Just before we go to you, we are in the middle of a series of committee discussions and we will come back in a second. But Stephen

CO-CHAIR GROSSBART: Yes, I wanted to comment and a couple of points have been raised, and I'm echoing what Peter said. The low percent performance among providers reflects the nature of this all in one measure.

And this is not an all in one process measure, but it's got outcome -- it includes outcomes. It actually is outcomes, except for the written action plan.

So these numbers are not unique and there are similar measures right now for diabetes care, as Peter noted, there's also

the D5 and a cardiac care which is the C4.

They are outcome measures, and I think the unintended consequences of providers looking bad in the community, I think that's not been an issue in markets where the diabetes 5 have been adopted.

We have adopted them in Cincinnati. They are used in Minnesota. And they are driving improvement in care and these are really things that patients care about and they're -- some of them are tough, tough to achieve, and but I think the philosophy of these measures is very, is very important for patients, and I'm sure that consumer groups will echo that, although I don't think we have a consumer representative on this committee.

CO-CHAIR WEISS: So, Christine and then Rubin and then -- Christine, did you raise your -- no. Okay. Don and then Dianne. Okay. So Christine, did you raise your hand for something?

MEMBER STEARNS: I was merely

noting that there are representatives.

CO-CHAIR WEISS: Oh I'm sorry.

MEMBER STEARNS: Sorry.

CO-CHAIR WEISS: I apologize. Okay great. Okay thanks. So then we'll go to Rubin, Don and then yes.

MEMBER COHEN: I'm just wondering, this asthma plan, is it standardized? Everybody has the same plan, they check off boxes? Or is it individualized to the clinic, to the patient?

CO-CHAIR WEISS: Yes, so we now have the measure developer on, so Rubin, if we want to, do we, again, now we are talking about the Minnesota measure, so the question is, Rubin?

MEMBER COHEN: So, if you have the asthma plan, is it -- you hand everybody the same plan and they check off boxes, how it suits the patient, or each clinic, each doctor comes up with their own plan with the individual patient?

CO-CHAIR WEISS: Did you hear that on the phone?

MS. PITZEN: Can you hear me okay?

CO-CHAIR WEISS: Yes, we can hear you just fine. And the question is, is if you can describe with a little more detail what it means to successfully complete the component of the asthma management plan.

MS. PITZEN: We are not requiring a standard asthma plan to be used by all clinics. We are requiring that the plans contain written components.

And those components are medications, dose and purpose, recognizing what to do during an exacerbation, and validation process against what was stated.

CO-CHAIR WEISS: Okay.

MEMBER GLOMB: What about the triggers that's stated here in the definition of that asthma action plan?

CO-CHAIR WEISS: And what about the triggers in the action plan?

MS. PITZEN: There's the expectation that those triggers be documented.

MEMBER GLOMB: Okay. So if they are not, they would fail that measure?

MS. PITZEN: So if one of those components is missing from that component, then that piece fails and then the measure would fail as well.

CO-CHAIR WEISS: Now, we asked some of these questions about the survey and survey response and you gave us the comments back. We are going to scroll to that section on our screen.

But did you want to talk a little bit about those additional findings that you had?

MS. PITZEN: Sure, and you know, I -- we missed the full discussion so I don't know if you wanted me to back up and describe the measure to you.

CO-CHAIR WEISS: Well no, we have gone through the -- in the absence of having

you here, as the primary reviewer being me, I also happen to be co-chair, Kevin Weiss, I walked them through that, and what you are getting are specific questions where they may have particular interests.

One of the issues of course was trying to better understand this asthma control questionnaire and because it was an important piece of this, and you had -- the working group asked for some more detail and you had actually worked to get us that, so could you talk a little bit about that?

MS. PITZEN: Happy to address. This is a fairly new measure released --

MS. BOSSLEY: Collette are you on speaker? Because if you are on speaker, you are breaking up and it may be better if you pick up the phone.

MS. PITZEN: Okay. Will I disconnect? Can you hear me okay now?

MS. BOSSLEY: I think you're fine. Go ahead.

MS. PITZEN: Okay, so we are using four validated asthma assessment tools and part of the all or none composite is that if that tool had not yet been implemented or used for that patient, they were counted as a numerator miss.

When I did an additional analysis, and this was implemented statewide, so really a really large population, when I looked at just the patients who had all three components as part of their medical record, then we were at a 63 percent achieving the optimal asthma care score, meaning they'd met all three components of the measure.

So we are fully anticipating that in our second year, second data cycle, that those actual optimal care rates will increase significantly.

If you can hear me, I can't hear anything.

CO-CHAIR WEISS: No, we weren't saying anything at that moment.

MS. PITZEN: Okay.

CO-CHAIR WEISS: We're just, we're soaking it in. Okay. Very good. So I think we are about ready to look at impact and then, Don before we do, we'll make sure you get a chance to chat.

We're getting close to where we can talk about impact, performance gap --

(Alarm sounds)

CO-CHAIR WEISS: So that's 15 minutes on this, and evidence. But let's do a few more questions then we'll vote on that and then we'll go to public comment, go to lunch and finish the measure after lunch.

But Dianne.

MEMBER JEWELL: I'm sorry I don't know the history of this measure. Were each of these individual performing measures before they were put into a composite? Were they tested and worked well as individual standalones, the three elements?

CO-CHAIR WEISS: Did you hear the

question?

MS. PITZEN: We were not able to hear the question.

CO-CHAIR WEISS: Okay, so I'll repeat the question, and that is was there testing on the measures individually before they were put into composite?

Now this is more of a validity question but as long as you are on the line let's do that.

MS. PITZEN: Sure. The measures were not tested individually per se before the composite was developed. However we do have the individual component -- that measure and we also publicly report those pieces of --

CO-CHAIR WEISS: Is there -- and I guess the question would be, is that there's evidence for each of the three components, but now the question is, would be, is the validity of those, and what we are hearing is that you are testing the validity of those as we, sort of as -- concurrently.

MS. PITZEN: Right.

CO-CHAIR WEISS: Good.

MS. PITZEN: This is very similar to our diabetes measure. Again, we are looking at all of those -- opportunity and to understand the measure better.

CO-CHAIR WEISS: Very good. Don.

MEMBER YEALY: So my concern had to do with with the outcome measurement. We are treating hospitalization and emergency department visits as essentially equal weights, and those are dramatically different events.

So two of either one of those gets you above a threshold or beneath one, and that simply lacks face validity to me. I mean, I think three ED visits in a year is a whole lot different than three hospitalizations.

If that's the major outcome, then they are not weighted at all, and I struggle with that.

CO-CHAIR WEISS: So that's a good

comment. We are into validity again, which is fine. So let's take a vote right now on the first element of this, and then let's go into public comment.

DR. WINKLER: I think we need to finish the measure. I think it will be too disruptive if we didn't.

CO-CHAIR WEISS: Then let's do the -- let's see if we can finish the measure then. Impact. So we want to look at the impact of this measure, as it relates to, is it an important specific national priority, or data has demonstrated high impact on healthcare improvement.

So one, two, three is high, moderate low, and four is insufficient.

(Pause for voting)

CO-CHAIR WEISS: Thirteen say high, six moderate, no low and one insufficient. Let's go on. This is the gap, that the data has demonstrated considerable variation and less than optimal performance.

(Pause for voting)

CO-CHAIR WEISS: One, two, three and then four is insufficient evidence. So we have 11 high, 5 moderate, 3 low and 4 insufficient. This time is was 4 insufficient. Zero low, oh, back to that again.

I was so excited about 4 insufficient that I -- next. Okay. And then evidence. Is there enough evidence to support the measure, yes, no or insufficient evidence.

(Pause for voting)

CO-CHAIR WEISS: Again, remember to press send. Almost there. There we go. Oops. Okay. So 5 say yes there is evidence, sufficient, 1 says no and 13 says insufficient evidence.

So to our colleagues in Minnesota, what we are hearing is at this point we cannot move the measure forward because of the NQF rules, because of this being a critical element.

What we have had is a lot of discussion, some of which you were able to participate in, some of which not. I think it's all recorded so there will be a way for you to understand what happened.

What you heard was a lot of interest in this measure, and a lot of support for the concept of a composite measure, and a lot of questions, a lot and a lot of questions.

And those questions, I think you are on the way to answering and gaining some of the evidence that will actually allow this committee over time to be real supportive of the direction you are going.

But at the current time, I think we are just left, from a committee vote, of saying we're not -- it's not yet ready to move forward on an evidence base.

Is that -- do I have that in terms of gestalting, what I heard from the committee? I'm seeing some yeses on this

side. I'm hearing, seeing some yeses on that side.

Okay. Good. Thank you so much for joining us. Okay good.

MS. PITZEN: This is Collette. I just have a question. We did miss entirely that whole discussion of the evidence, do if we could get reporting or some minutes or something so that we would have some direction.

DR. WINKLER: Collette this is Reva. We will have the transcript for you next week, and we can show you that, that's not a problem.

MS. PITZEN: Great. Thank you very much.

CO-CHAIR WEISS: Thank you.

DR. VIR: This is Dr. Bani Vir here again from ActiveHealth. You know, we have been waiting for quite a while to defend our measure, and because of the technical difficulties today, we understand that you are

-- that you -- the plan right now is to break for lunch.

But we have had, you know, we have a group of clinicians here who have taken time out of their busy schedules to be at this meeting, and we were anticipating completing this measure by 1 p.m.

We would really appreciate it to have the opportunity to defend the measure now, before you took your break, to allow our clinicians to also return to their days.

CO-CHAIR WEISS: You anticipated us by about 15, 20 seconds. So we recognize that the group here is ready for lunch, however we also do respect the fact that we have had some technical difficulties.

So I was going to ask the group just for a moment, would it be okay, for the group to do a short but definitive revisit of our -- of the measure that was presented to us by ActiveHealth, just so that they know what happened in terms of our thinking and then

give them a chance to comment and see if that would lead to any additional discussion on our end? Are we coercing you all into holding off for food for a few minutes? Let's do it. Okay good.

So why don't we start by just making sure that we hear the measure as you would like us to have heard it. So if you can give us like a one minute to two minute look at your measure as you see it, and then we will give a sense of how we have seen it to this point and then the issues that we have raised and give you a chance to talk about those.

DR. VIR: Sure, so as I mentioned earlier, the measure is really looking for the percentage of people who have had access to at least one rescue inhaler in the past 12 months.

Again, the spirit of this measure is not to look -- is not to delve into long term control, appropriate use of appropriate

controller medications. It's really to have a rescue inhaler available for emergency use.

And I understand that there were some concerns regarding things like capturing the appropriate population, as well as ample use and some other issues.

So just to clarify some of those issues, first of all, in regards to capturing the appropriate population, we take great care to make sure that this measure is highly specific. Our denominator doesn't just look for an asthma diagnosis any time in the past.

We look specifically in the past year for multiple diagnoses overlapping with office visits, overlapping with asthma medications that are not short term, that are not rescue inhalers to confirm that the patient is truly asthmatic.

Also a lot of our patient and provider feedback is telling us that the patient truly doesn't have asthma, and if they do give us that feedback, we pull them out of

the denominator --

When it comes to samples, we allow for patients and providers to also tell us that they have given the patient samples. The shelf life for the medication is addressed.

We understand that the shelf life of albuterol is two years. I know there was some concern about that. However, we have no way of knowing when the medication was dispensed from the time it has arrived on the pharmacist's shelf to when the patient actually received it. We don't know what that time gap is.

So we have to -- we decided as a collective team to look back one year because the efficacy, as you all know, of albuterol, decreases over time, and we didn't want to admit people and in term have our measure, you know, give it erroneous results that could allow for increased hospitalizations and ER use.

CO-CHAIR WEISS: Thank you so much

for that input as well as a particular focus, because you did touch on some of the concern of the committee.

I don't know, Rubin, I am going to ask you if you are -- this is kind of your heads up moment -- did you want to reflect on how you heard the committee's response to the review of the workgroup or your thoughts on this measure, or do you want me to go ahead?

I thought you would say that. Good. Okay. So the committee, in thinking through this measure, heard the concerns, the principle concerns about the potential for mis-classification, at least that's the way that I would phrase it, in the -- and because of that, whether the gap was quite what was we were seeing.

So when we saw the gap that was presented to us, we said that seemed to be awfully big, and it opened up the big question about, about the measure specification in terms of these issues that you've raised, and

the evidence supporting that just having a dispensing is enough to get us there.

The -- as we went further, looking at the gap, we were really unclear about whether we believe the gap that we were seeing, because of the fact of the shelf life, and because of the uncertainty about the sampling, the sample process, as well as the denominator being such a broad net for asthma, that in fact it may not require a dispensing in all cases.

Again, a lot of uncertainty there. So without a real good sense of certainty of the measure, in terms of validity, we didn't know that we had good information on the gap, and so it did, on voting, it came up as having a high degree of committee members who were feeling like they had insufficient information to make a decision on this measure right now.

What I sense that the committee would like, and we didn't talk about this formally, would be, is if one was to look at

this measure again in the future, would be to really understand the nature of that 40 or so percent -- actually almost 50 percent who are not getting these inhalers, and wondering, are they really the people who should have been getting them that were not, or were they people who had them on the shelf and were not, and to sample into that population so that we really understood that there was a performance gap that had to be fixed.

Once that was done, that would also probably clarify a lot of the validity issues that would come later in a discussion that the committee did not have because of the gap was where we stopped.

Do I have that correct? I have --

DR. VIR: I'm a little concerned about --

DR. WINKLER: Bani, hold on a sec. I'm going to take over for Kevin while he --

DR. VIR: Can I say something?

CO-CHAIR WEISS: Oh excuse me one

second. Reva would like to add another additional comment on reflection of the discussion.

DR. WINKLER: Yes, Kevin is coughing and drinking water to clear it up. Just in terms of the gap, there were a significant number of folks who had concerns about having insufficient information to really understand that number.

However, the real I think telling point was under the evidence criteria. The concerns were registered about the overly large denominator and that the construct around the idea that the absence of a medication dispensed may in fact not represent a need for that medication, and the evidence that supports that is not clear.

And within your submission, you do note that there are no major studies formally assessing the absence of rescue therapy in asthmatics and there's very little published data regarding asthmatics in the presence of

short acting beta agonist to prevent asthma attacks.

So when it came down to the vote for evidence, the majority of the committee voted that it was insufficient to support the focus of the measure. So that's just a summary of what happened.

DR. VIR: Okay. So I'd like to express a concern that I have because it seems that you all have voted when we were obviously off the phone, and we weren't connected, and you came to a conclusion without hearing our explanation and made that vote without -- before hearing the explanation.

If you'll let us clarify the -- how specific denominator is, and it's quite specific. I mean you would be hard-pressed to say that the people we are capturing are not true asthmatics.

I think I explained already that we are looking not just for a diagnosis, we are looking for diagnoses overlapping with

office visits with the same diagnosis code overlapping with asthma medications that are not short acting but long acting asthma medications, as well as provider and patient feedback.

I don't know how much more accurate you can be in identifying a true asthmatic, and we only look back in the past year to prevent that sort of diagnosis carried forward from an old chart or from two or three years ago.

Additionally, the gap that you are talking about, you know, we take in data from every possible source that there is, whether it's a health information exchange, pharmacy data, administrative claims, patient, provider feedback. We have patients talking to our nurses through telephonic engagement and disease management programs. Providers tell us whether or not they have -- we have got the diagnosis correct.

So to say that 42 percent is not a

true gap, I mean, we are talking about electronic measures here and we are obviously limited to what is captured electronically for the for -- other developers are limited to that.

However we take it a step further, actually many steps further, and allow for feedback to be given to us and entered manually.

So I'm not sure what -- where the concern is here. It's really not clear tome. As far as the literature piece goes, I think that it would be highly unethical to actually conduct a study where you withhold short acting beta agonists and those studies would be hard to find.

I'm going to also defer to our subject matter expert, Dr. Sharma, just on the literature piece and you know, a deeper guide.

DR. SHARMA: Yes, I mean my only comment would be that you know, it's very difficult to find studies on the literature

that are looking at what's basically the cornerstone of asthma therapy.

Now, given that, I mean, we have -- this is a sample set of like 385,000 members with age ranges between 12 and 77. Now, looking to Dr. Bani Vir's point, looking at the fact that we are being that specific, so that the asthma code overlapping with office visits, plus some months of medication.

And with that said, I think that's fairly sensitive and specific, the denominator, to Dr. Bani Vir's point, is we are taking patient and provider feedback that would remove you from the denominator, and the fact that we are finding a gap, to me, isn't actually surprising, because we, you know, I'm in internal medicine and when we treat asthmatic patients, you are more worried about the problem of inhaled corticosteroid or long acting beta agonist and that sort of short acting rescue therapy sometimes falls through the cracks, and that's exactly what we are

seeing.

With that said, we are looking back a year, we stopped like maybe two, to Dr. Bani Vir's point, we don't know when the time period was from when it was made at the factory to the pharmacy and got to the patient.

And so I think there is here, there is a true gap here, I mean, 52 percent compliance, which is surprising because no one has looked at it, and we have actually taken the time to look at the fact that are there members in the population, or patients in the population, that are not -- do not have access to short acting beta agonist therapy in the past year.

And I think we should actually approve this measure and we can show followup data next year to show that this is in fact true.

CO-CHAIR WEISS: So thank you. That's been very helpful. I must dust off a

little bit of my knowledge of the asthma literature, and I know we have got people here who probably are contemporary with it more than I am, but I think there's a number of studies from emergency rooms who have done intake audits about what medicines people are on, and I don't know that there's an absence of -- at least unless it's a newly-diagnosed asthma coming into the emergency room -- for people who have asthma, that they actually come in with beta agonists. The big problem is that they are not coming in with long -- with anti-inflammatories into the emergency room, not with the beta agonist.

So I think it would be nice for you to pull that literature and make sure that that's consistent with what your findings are.

But that aside, I think it's really good what you have given us a chance to present -- what I'd like to do as a committee is to see whether or not you would like to reconsider our action.

I think it's a very fair question being that we had not had the very good input that we just got from the measure developer.

So if you are interested in opening up the measure for re-discussion, we should do that, and let me get a sense of the table as to where you would like to go with this.

So what I'm going to do is just ask for a yes, no, and that is the yes would be yes, we will open up the measure for re-discussion based upon what we have heard. No would be is you feel like that although it's been helpful to hear this, that we don't feel like we want to open up the measure again for our discussion at this point in time.

Yes or no, if that's okay. Does that work from a staff perspective? Are we okay to do that? Okay great. So let's see how many people would like us to reopen the measure, please raise your hand.

(Show of hands)

CO-CHAIR WEISS: Okay, and how many would like us to, at this time don't feel the need to open it so no. Raise your hands.

(Show of hands)

CO-CHAIR WEISS: So we have -- so for you on the phone, it's just -- we only had one member who was feeling the need to reopen the measure. The rest, 18, next to me, 19, I don't know what your -- overwhelmingly in favor at least right now, of staying where we were.

But we want to thank you for holding on with us and coming back and taking some time out and presenting.

DR. SHARMA: I'm sorry, this is Dr. Sharma. Just one more comment. You know, you're looking at -- so the comment about everyone in the ER having a short acting beta agonist. So that's sort of after the point, right?

I mean they're already in the emergency room. But what we are finding are

people that have not been in the emergency room yet, so you are looking at a subset that is not being well controlled, that may have the short acting beta agonist. We are talking about being more preventive, preemptive, to say let's try to prevent that ER visit because we don't see any claims evidence for the short acting beta agonist.

So I mean, I hear the comments about people, asthmatics in the ER having a short acting beta agonist. We are identifying people that don't even have a refill within a year, for a short acting beta agonist.

So I mean, I do with all due respect actually disagree with the committee. I think, you know, we are missing on a very cheap point, if we are looking at outcomes and trying to prevent ER visits and hospitalizations and keep costs down, and we are assuming that every asthmatic has a rescue inhaler therapy, and we are telling you, given our data set, you don't even see it in 48

percent of the population.

With that said, you know, I, you know, I will, you know -- I will agree with what the committee says, but I think that we're making a very poor choice here in not approving this measure, because we are identifying people that don't have a short acting beta agonist therapy.

CO-CHAIR WEISS: Well received, there is an opportunity for comment back to the committee, and I think the other important point is, is that the issue was -- that you heard from the committee was insufficient evidence on these issues.

So as you collect more evidence that support your measure, it will add weight to a reconsideration at some point in the future I would think.

But for the time being, it was not a no because we didn't agree. It was a no because we didn't feel we had enough information based upon the concerns, and I

think you should take that as a formative bit of feedback that was positive from the committee.

But we have to close this discussion because I know we have to go to public comment.

So I want to thank you very much for your input and I think we need to now move into public comment mode. Anyone in the room, away from the table who are public here, want to make comments?

(No response)

DR. WINKLER: Anybody else on the phone?

OPERATOR: For public comment over the phone, please press *1.

(No response)

CO-CHAIR WEISS: Thank you all for -- on the telephone for your participation for the committee. We have lunch. Can we compress it maybe about 10 minutes and maybe aim for a 20-minute lunch, and if you want to

come back to the table with food so that we can eat a little bit as we start up, that would allow us to continue and a quasi-working lunch.

So about 20 minutes from now, we'll start back up and welcome you to the table with food.

DR. WINKLER: Restart at 1:15.

(Whereupon, the proceedings in the foregoing matter went into lunch recess at 12:56 p.m. and resumed at 1:20 p.m.)

A-F-T-E-R-N-O-O-N S-E-S-S-I-O-N

(1:20 p.m.)

CO-CHAIR WEISS: So let me start our after-lunch version of a continuous evaluation process. Are you all ready to begin, yes? Good.

We have -- we are shifting from asthma to pneumonia, and no, we are -- before we do pneumonia we have a couple actually of measures in before.

We have Measure 0513, which is the CT of the thorax measure, and do we have Dr. Kazerooni with us?

MEMBER KAZEROONI: Yes, I'm on the phone. Thank you.

CO-CHAIR WEISS: So that's wonderful. Have you been with us at all today, otherwise, just so I can know whether or not we need to get you up to speed in terms of process?

MEMBER KAZEROONI: I'm just joining.

CO-CHAIR WEISS: Okay very good. So first, welcome, what you have got is Kevin Weiss who is one of the co-chairs and --

CO-CHAIR GROSSBART: Steve Grossbart is another co-chair.

CO-CHAIR WEISS: And we have the -- both staff here and about, I'd say, 16 of us around the table, plus or minus a few, and then we have some other folks here who have joined us in person as part of a more general public interest.

And what we are going to do is we are going to walk through the measure. The way we manage these is first to do an overview, and ask our measure developers, if they are with us, to give us a one or two minute, and then we'll ask you, Dr. Kazerooni, to give us in -- your review in sections.

And the first section we'll ask you to give on, has to do with the importance, the performance gap and the evidence. So we'll take those three items together, and if

you can sort of put your mind around those, that's where we'll be beginning.

So do we have the measure developers with us? I know CMS is the official measure host, but do we have a contractor with us?

RICH MAY: Rich May here.

CO-CHAIR WEISS: Rich? Do I have that right? Rich, are you there? Oh, which measure, sorry. 0513, thorax CT, use of contrast material.

RICH MAY: I can't speak to that one.

CO-CHAIR WEISS: Next okay, well then I guess Dr. Kazerooni then, would you be willing to give us a general overview, us being our committee?

MEMBER KAZEROONI: Certainly the measure regarding thoracic CT and the use of contrast material is something that has been the subject of a lot of discussion in the radiology community and the appropriateness --

committees that I work on through the American College of Radiology.

We are revising most of our published criteria to now specifically taste not just CT -- but whether it's with, with and without, or without contrast, and there is almost no circumstance under which we are recommending with and without contrast, so I think this is a very appropriate measure.

CO-CHAIR WEISS: Can you give us a view of the measure itself?

MEMBER KAZEROONI: Just a review of what the measure itself is?

CO-CHAIR WEISS: Yes, as part of the general introduction.

MEMBER KAZEROONI: Okay. CTs of the chest are very -- very infrequently would require them to be performed both with and without contrast, and the measure has basically used the total number of CT studies performed as the denominator, and the numerator to be CTs of the chest with and

without contrast.

This percentage would be a very low number in most practices who are performing appropriate CT imaging.

CO-CHAIR WEISS: Great. And with that as background, what we'd like to do now is start talking about the impact, the performance gap and the evidence. So can you review that both from your perspective as the reviewer, and then we'll ask the working group for additional comments to follow?

MEMBER KAZEROONI: I would say that there is definitely a range of performance if we were to apply this currently to practices today. The reporting of this measure already on a publicly available website has already had impact in reducing the frequency with which these combined contrast and non-contrast studies are performed.

I believe many people did not -- basically have not simply gone to the point of reviewing their protocols and the documents

reflect -- fitness, and we have seen very quick changes in practices once they have seen a Hospital Compare of their performance metrics relative to peers.

So I think there is a large gap in practice with respect to adherence to this. I think there's the potential to have a large impact and I think it should happen fairly quickly.

CO-CHAIR WEISS: So, in terms of the performance gap, what do we know about the performance gap, and I think we have a -- Reva was telling me we've got a screenshot of Hospital Compare.

DR. WINKLER: This measure is reported on Hospital Compare for hospital outpatient imaging facilities and the national average is 0.05 and -- oh shoot it's too small -- it's the third of the measures they grouped together so the chest thorax is the third one. You can see that the national average is 0.05. And then I picked three random hospitals in

the local area, and it really does range from 0.01 to 0.05, so there is variation.

And so I'm in -- I believe, I believe, interpreting the way they say that it will be five percent, or one percent, or -- and I don't know why it's portrayed as the decimal as opposed to some of the others, which are the percent.

CO-CHAIR WEISS: So we're seeing between a one and five percent variability. And do we know anything about trending in this -- for this measure?

MEMBER KAZEROONI: I don't have any formal information about trending. All I can say is what I'm aware of in individual practice circumstances, where as soon as they have seen their information on Hospital Compare, they have immediately addressed it in their practices as being outliers.

I think most of them are not even aware of this.

CO-CHAIR WEISS: And to those of

us who don't know this area well, when you say address it, how far off will they be from this one to five percent that they would come into line, would you think?

MEMBER KAZEROONI: Places that have come into line, where we would consider it to the level of appropriateness, well under one percent of chest CTs should be performed in this manner.

CO-CHAIR WEISS: Under one percent is what we heard. Okay.

MEMBER KAZEROONI: Pardon me, can you repeat that question?

CO-CHAIR WEISS: That's Peter. You're -- Peter you need to make sure that your -- there you go.

MEMBER ALMENOFF: I was just trying to figure out what the percent is. Is it five percent right now?

DR. WINKLER: Dr. Kazerooni on that Hospital Compare, the reported national average is 0.052. Are we interpreting that as

5.2 percent? Is that correct?

MEMBER KAZEROONI: That's my understanding.

MEMBER EDELMAN: I'm sorry. I'm not understanding what the measure is. The visual says combination scan. Is that two scans, a plain scan followed by a contrast scan?

CO-CHAIR WEISS: Yes.

MEMBER EDELMAN: So this does not include a planned CT with contrast?

CO-CHAIR WEISS: No.

MEMBER EDELMAN: This is only for that practice of a plain scan followed by a contrast scan? Thank you.

MEMBER KAZEROONI: Yes, but there's a specific CPT code for chest CT with and without contrast. There's one for with alone. There's one for without alone. And it's that combined with and without contrast in the same setting that is really inappropriate in almost all -- in all

circumstances.

There are other chest CT billing codes to be aware of that are CT angiographic codes. They are CTA codes. Those include with and without contrast as part of an angiographic study, and one of the things that we have seen is that people are miscoding some of their exams and should be using a CTA code instead of using chest with and without codes.

MEMBER EDELMAN: So this does not include the CTA?

MEMBER KAZEROONI: This is not -- this measure does not include CTA at all, because they are separate codes. And those codes include with and without within them.

CO-CHAIR WEISS: Mitchell?

MEMBER LEVY: So if I understand correctly we are looking at a negative performance metric, where we are measuring the amount of time this is being done inappropriately? It's unusual.

MEMBER RHEW: You had mentioned

that, first off, it should never be done, but then you said, well, it should usually never be done. But I just want to clarify. Is there ever a circumstance that you would ever give with and without, you know, for any -- I'm just trying to figure out, is there anything that goes into the exclusion category?

MEMBER KAZEROONI: There are some very narrow indications for performing a CT with and without contrast in the same setting. These are things such as CT tumor perfusion studies which are performed in very high academic medical centers, doing things like radiofrequency and cryoablations for lung cancer, and are not really mainstream practice.

So most practices should almost never be billing this code.

MEMBER ALMENOFF: So does Medicare pay for this code right now? Does Medicare pay for this code right now?

MEMBER KAZEROONI: You are breaking up, can you repeat the comment?

MEMBER ALMENOFF: I was asking if Medicare is paying for this code right now.

MEMBER KAZEROONI: I got the last part. What was the first part of what you said?

MEMBER ALMENOFF: Is Medicare paying for this code, this before and after code right now?

MEMBER KAZEROONI: Yes they are.

MEMBER ALMENOFF: And if it's not practice or it shouldn't be done, why would you even fund that code? I mean it's all about money --

MEMBER KAZEROONI: There may be narrow circumstances in which it may be appropriate.

MEMBER ALMENOFF: So wouldn't there be like an exception rule where you can ask for some additional resources but I mean, wouldn't that be an easy way to just eliminate

this by stop paying for it, and stop using a performance measure to do this?

MEMBER KAZEROONI: I think there still is the need for this code because of some of the narrow clinical circumstances under which it is performed, usually related to tumor imaging.

But it is a very narrow, clinical indication to do this. So it's not zero, but it's very small.

CO-CHAIR WEISS: So, Norman?

MEMBER EDELMAN: I understand the economic interest but I don't understand the health outcome. I mean my guess is, certainly knowing the radiologists I know, that all you do is take that five percent and convert them all to contrast studies.

So what is the evidence, even theoretically that this is going to have any impact on health outcome?

MEMBER KAZEROONI: If they're actually performing a CT study with and

without contrast, then they are exposing the patient one, to unnecessary radiation, and that has its potential downstream consequences in exposure to radiation, with the possibility of increased risk of cancer, so that's a very real health outcome, hard to track in an individual patient but believed to exist on a population basis.

There are some of these cases that do not require contrast and I really don't believe they would all be converted to de facto with contrast examinations.

Some of these would become without contrast examinations. Some would become with contrast. And some would actually become CT angiographic codes.

MEMBER EDELMAN: I believe most of them would become with contrast studies. I'm not sure you are going to get the outcome you want.

CO-CHAIR WEISS: As I imagine that hasn't been looked at directly with those who

have improved this negative measure to see what's happening to these folks but that would be of interest.

Let's be mindful. I think we have had a good discussion on impact and gap, and on evidence. Mitchell?

MEMBER LEVY: I agree, but I think we need clarification about this 0.052, because it also could be less than one percent. So that's important. If it really is 5.2 percent, that's very different than 0.052.

CO-CHAIR GROSSBART: I just pulled up Hospital Compare while we are talking. It's measured on a zero to one range. So 0.05 would be five percent.

CO-CHAIR WEISS: So, in terms of impact, it doesn't seem like it's a lot of people, but it is potentially related to a theoretical outcome, and I guess there is, for those who would be not necessarily getting contrast, that are getting it now, there's a

theoretical possibility of dye reaction or some -- I mean, but these are small numbers in terms of impact in that sense, not necessarily because of the scale of how many people are getting them.

MEMBER KAZEROONI: So I guess I would look at the outcomes as being reduction in radiation exposure, a small reduction in contrast dosage administration, and then a reduction in charges and cost.

CO-CHAIR WEISS: Okay. Very good. So let's now go to voting. Is everyone okay to go to voting? Okay? Oh, and just a note, I've -- sorry, but I lost my process -- did everyone in the workgroup get a chance to speak to what they thought about this?

(No response)

CO-CHAIR WEISS: Okay. And then we had the committee as a whole. Let's go for a vote. Impact, high, moderate, low and four is insufficient.

(Pause for voting)

DR. WINKLER: Dr. Kazerooni how would you vote on the rating for impact, high, moderate, low, insufficient?

MEMBER KAZEROONI: I would vote for moderate but I don't yet know how to triangulate my response to how other measures are reported. But I might as well just start.

DR. WINKLER: Okay.

CO-CHAIR WEISS: Make sure you're pressing the button. There we go. We got 20 here plus Dr. Kazerooni. Okay, so 3 said high, I'm going to make it 11 moderate because of Dr. Kazerooni, and then 7 low and none insufficient.

Good. Next. This is the gap question, and let's vote one, two, three, high, moderate, low and then four would be insufficient.

(Pause for voting)

MEMBER KAZEROONI: Moderate.

CO-CHAIR WEISS: Okay. All the numbers coming in? Let's give it another hit

of the number everybody. Okay. So let's do it again with everybody pointing to Jessica.

Okay, let's turn it around three times, again. One more time, again, you just have to try and get -- one of these may have just a bad battery or maybe something going on. We had it just a moment ago. For the -- we had a vote just a moment ago with 20, so we're all here. It's hanging on there, well, it's going to show us anyway with 19, right, because it timed out. Okay that's fine.

So, 3 high -- that's 20. So, 3 high, we're going to make it 11 moderate, 7 low and no insufficient. So it passes all three characteristics.

No, sorry, two of the three. Evidence. Is there sufficient evidence? Yes, no, or insufficient evidence.

Yes, no.

MEMBER KAZEROONI: Yes.

CO-CHAIR WEISS: Okay, and you're saying yes, okay.

(Pause for voting)

CO-CHAIR WEISS: There it goes. Okay. So, 16 yes, 1 no and 4 insufficient. So it passes this -- we go into reliability and validity. So if you could now present to us your thoughts on reliability and validity on the measure?

MEMBER KAZEROONI: Could you give me a little background on how you usually describe this?

DR. WINKLER: Essentially the measure evaluation criteria is that the measure has had testing of it, either at the level of the data element or at the level of the measure score, or optimally, both, to determine whether the elements or the results are reproducible and reliable.

Validity on the other hand is, given the result that's generated from the measure, do -- is it an accurate reflection of quality that can be demonstrated through empiric testing or commonly will see face

validity.

MEMBER KAZEROONI: Well, I guess reading from the documents that have been circulated about this measure, this is believed to be reliably reported and the dry run that is described for Medicare at over 3,000 hospitals have downloaded their specific information; during the dry run processing, note that over 500 emails were submitted with questions about this particular efficiency measure, and there are very few comments that were received about the chest CT one specifically.

Their conclusion was that the low level of inquiries about the specification of this measure, that they inferred that the results are reliable and so it's a fairly straightforward metric to collect the information for.

CO-CHAIR WEISS: Great. So your -- we are hearing a high degree of reliability and validity and probably because it is based

upon billing data.

It looks like from what we see here that the workgroup also rated it high in reliability and validity. So additional comments from other members of the workgroup on reliability and validity?

(No response)

CO-CHAIR WEISS: Okay. Any questions from the committee at large on reliability and validity? If not, let's vote.

(No response)

CO-CHAIR WEISS: Okay, let's vote then.

MEMBER KAZEROONI: I would vote high.

CO-CHAIR WEISS: Okay, thank you.

(Pause for voting)

CO-CHAIR WEISS: Got it. Let's see what we've got. We have 16 high, 4 moderate and no low and no insufficient for reliability. Let's go to, yes, to validity.

Again, let's vote high, moderate,

low for validity, and four for insufficient.

MEMBER KAZEROONI: High.

CO-CHAIR WEISS: Okay.

(Pause for voting)

CO-CHAIR WEISS: Let's vote -- repeat our vote please everybody. There we go. Done, 14 high, 6 moderate, no low and 1 insufficient. Good. We go on to the final sections of usability and feasibility.

So let's talk about usability. Is the measure meaningful, understandable and useful for public reporting? And whether it's meaningful and useful for quality improvement.

So, Dr. Kazerooni, any thoughts here?

MEMBER KAZEROONI: I think this is relatively straightforward, easily understood and meaningful in terms of public reporting and understanding of this metric.

CO-CHAIR WEISS: Okay and that's reflected also in the workgroup having predominantly high usability and feasibility

ratings. So from the members of the workgroup, comments or thoughts before we go to a general committee?

(No response)

CO-CHAIR WEISS: Okay, general committee then? Dianne.

MEMBER JEWELL: So given the question that I think Mitchell asked earlier, there's not a concern that people, anybody would misunderstand that this is actually looking for a -- this is a negative indicator, if you will?

And I know we don't have the contractor on the phone, right, so we don't --

DR. BROOTMAN: I'm sorry, this is Dr. Brootman, I'm a contractor who developed the measure.

CO-CHAIR WEISS: Very good. Dr. Brootman, what might be your response to the confusion of the, of the end user on this one? Or consumer I guess.

DR. BROOTMAN: As the consumer,

well, this is, I would say what -- probably the consumer doesn't have too much say on the decision on with and without contrast.

You know, that's basically one, and compared to others where they can decide on this is something that is decided completely by the decision at the time of doing the study.

Now I think there is information that the consumer would want to know regarding, and can make a distinction, if they understand -- you know, on the risk of having contrast when it's not necessary and additional radiation which is not necessary, and doing an additional, what's called an additional study that -- and that's why the meaningful -- the meaning of public reporting is very helpful for patients to at least acknowledge and make a decision on the studies.

Obviously the decision on getting or not with and without contrast is not going

to be in their state of mind, but I also, you know, there's been a lot of public awareness on this, beyond the public reporting, this was a number of very well -- very well-known articles in the New York Times and Washington Post --

(Alarm sounds)

DR. BROOTMAN: -- regarding the use of with and without contrast on CT, thorax and a decrease in the recent years because there is more and more evidence that there is no need for this double study, and -- or combined study.

And I just want to clarify, there are no stated exclusions for this. There is no evidence provided that there was any need for the double study or combined study, so I hope that addresses your question.

CO-CHAIR WEISS: Does that answer your question? Yes, I think what we were concerned about, since this is Hospital Compare, that it goes to the public, and I

think what Dianne was talking about was that even though it's technically something that is going to be acted on by the radiologist and the physician community more generally, it will be something that, if a person looks at Hospital Compare, they'll say whoa, my institution has a really low rate of this, and not quite understand why they are doing so poorly. Is that what I was hearing from Dan?

MEMBER KAZEROONI: This is something we would encourage a patient who is coming for any CT examination to be aware of, and to ask, usually it's the technologist who interacts with the patient around this examination, to ask how their scan is going to be performed, will it be a quote, double scan, as they have come to be termed, or not.

DR. BROOTMAN: I think it's important, you know, putting in a contrast substance, I know they've improved but you know, if you can avoid that when there's no need, I think it's a good question for

patients if they recognize that there's no benefit from doing a combined study, the need -- it would be a question for the patient to ask what are the benefits.

So I think it does help in making an informed decision for patients.

MEMBER RHEW: I don't think it's a question of the quality metric for the validity. It's really how it's presented. So you know, on Hospital Compare, they could have some of it says good care here on the left, and bad care on the right, they just flip it around for this. But it's not the metric. It's just how they present it on whatever site, so I don't know if that had any bearing on the metric itself.

MEMBER LEVY: But I think this is the metric, because I mean I'm not aware of any metrics that are negative like this. So if the metrics are for the benefit of the public, especially of public reporting, if we have one metric that's negative, it will be

almost impossible for the public to really understand that.

Because when I first saw that, I also thought boy, the compliance is really low with this metric. So it feels to me we are setting ourselves up for failure with this.

CO-CHAIR WEISS: Is it just, and maybe it's perhaps in the name. There are other overuse measures that you want to have low, and -- but it doesn't specify itself as saying overuse of, and if it had that in the title, that would probably be helpful. So that might be just a comment and a guidance statement back to staff that we are, as a committee, seeing a number of individuals concerned with the interpretation of this and that's a reflection.

But let's not stop there. Let's make sure we have got the usability issue fully covered. We've identified this issue of the naming and the perception.

MEMBER RHEW: The very simple

solution is you just call it appropriateness and you just do one minus, and it's just, you know, you just flip it around.

So I mean, I think that could solve the whole problem.

CO-CHAIR WEISS: Good. That's great. Stephen?

CO-CHAIR GROSSBART: Just a quick comment again on Hospital Compare. It specifically says that a high number may indicate overuse or too many patients -- to quote, a number close to one may mean that too many patients are being given a double scan when a single scan is all they need.

CO-CHAIR WEISS: Great, so let's go to a vote of usability. One is high, two is moderate, three is low, four is insufficient information.

MEMBER KAZEROONI: High.

CO-CHAIR WEISS: Okay, got that.

(Pause for voting)

CO-CHAIR WEISS: Oh, everybody

press again. Okay third time is a charm. Everyone point to Jessica. There it goes. Okay. So 10 high, 11 moderate, no low and no insufficient information.

Let's go to the final criteria here, which is --

DR. BROOTMAN: Can you repeat the numbers? We couldn't really hear you.

DR. WINKLER: It's 10 high, 11 moderate, zero low and zero insufficient.

DR. BROOTMAN: Thank you very much.

CO-CHAIR WEISS: For feasibility, Dr. Kazerooni?

MEMBER KAZEROONI: This is very feasible, this is all coded billing data, there are separate codes for with contrast, without contrast and with and without contrast. Some of -- there's the potential to have errors in coding given that there are three but that should be low, and if that is one of the errors it can be fixed, other than

that potential miscoding that an institution might be doing, this should be a very feasible, straightforward measure.

CO-CHAIR WEISS: Very good. And that's what the workgroup agreed with. Any comments from the workgroup? Any comments from the committee as a whole?

(No response)

CO-CHAIR WEISS: Then let's vote on feasibility. One is high, two is moderate, three is low, four is insufficient information.

MEMBER KAZEROONI: High.

CO-CHAIR WEISS: Okay.

(Pause for voting)

CO-CHAIR WEISS: Okay everybody, press again please. There it is, okay. So we're at 18 high, 3 moderate, no low, no insufficient information.

Let's go to the final vote for this measure, which is overall suitability for endorsement, that's a yes or no, one is yes,

two is no.

MEMBER KAZEROONI: Yes.

CO-CHAIR WEISS: Okay.

(Pause for voting)

CO-CHAIR WEISS: Okay everybody, press again. There we go. Good. So we have 21 yes. No nos. Unanimous decision. Great. Thank you so much, Dr. Kazerooni. We are going to do now -- we are going to shift gears again because Christine is going to be leaving. Stephen.

CO-CHAIR GROSSBART: So we are now going to shift to Measure 0179, improvement in dyspnea.

DR. BROOTMAN: Thank you so much. If you have any other questions, I'll be here. This is Dr. Brootman. Thank you.

DR. WINKLER: Thank you very much.

CO-CHAIR GROSSBART: And first thing is do we have a --

DR. BURSTIN: Shortness of breath works well too.

CO-CHAIR GROSSBART: Shortness of breath. Do we have the measure developer to give us a no more than two minute overview of this measure?

MS. DEITZ: Yes, Deborah Deitz is here from Abt Associates.

CO-CHAIR GROSSBART: Hello Deborah, Steve Grossbart here. Go ahead with your overview.

MS. DEITZ: Right. So as many of you know, CMS has developed a quality improvement monitoring system for home health over the past 10 years. It uses data that's collected via the OASIS data set, which is integrated into the home health clinical assessment.

That OASIS is collected for all the adult, non-maternity, Medicare and Medicaid patients that are receiving skilled home health services.

So this measure reports the percentage of home health episodes of care

during which the patient became less short of breath or dyspneic.

It's calculated on OASIS data that is collected as part of the patient assessment at admission and discharge. At each time point, based on patient observation, the clinician identifies the level of exertion that results in a patient's dyspnea or shortness of breath, using five behaviorally benchmarked responses that represent an ordinal scale.

CMS and the developers have a lot of confidence in this dyspnea measure and the data is based on, it was developed for the initial version of the OASIS in 1994, based on literature review, field testing, clinical panel input and demonstration pilot testing.

The item stem and response options have remained unchanged since their development, and more than 10 years of OASIS use by more than 10,000 agencies has found no significant flaws in the item.

There are three inter-rater reliability studies reporting percent agreement at the level of 0.82 and weighted kappa value of 0.49, and 0.51.

The outcome measure has been reported to home health agencies through the CMS quality website since 2001, and has been publicly reported on home health comparisons 2003.

It's been NQF endorsed since 2005. The measure specifications basically remain unchanged since their initial development, except that it now includes long term episodes since 2008, which was by NQF recommendation.

The measure is risk-adjusted using a very robust prediction model that includes 83 risk factors and has an R squared of 0.117 and a C statistic of 0.703.

For comparison purposes, most of the outcome measures based on the MDS have a C statistic in the 0.6 range, so it's very well risk-adjusted.

CMS believes that the improvement in dyspnea measure is important to continue to report for three reasons. Basically, dyspnea affects a large number of home health patients. It's an important health status indicator. It impacts qualify of life, can substantially affect a patient's ability to engage in a wide variety of activities, has been identified as a risk factor for hospitalization among Medicare home health patients.

It's frequently associated among home health patients with general deconditioning such as what occurs following a hospital stay with extended bed rest.

OASIS data indicate that 70 percent of home healthcare patients are reported as having some dyspnea interfering with activity.

The second reason CMS believes it's important to continue to report is that it's actionable. There are interventions that

can be implemented by home health agencies, that can improve dyspnea in many patients like teaching of activity pacing, problems with breathing, reinforcement of smoking cessation, and the correct use of medication.

CO-CHAIR GROSSBART: And may I ask you to wrap it up in about 15 seconds?

MS. DEITZ: Okay. The measure is used by a lot of agencies as part of a best practice improvement package, and it provides them with a data-driven basis for their quality improvement activities.

And the third reason of course is that the measure is important to provide valuable information to consumers via the CMS Home Health Compare website. That's it.

CO-CHAIR GROSSBART: Thank you very much. Christine, I'd first ask you to give a kind of high level overview of the workgroup and then let's go into the components of the voting.

MEMBER STEARNS: Well, after that

introduction I don't have a lot more to add. I would say for the workgroup discussion there were a few questions about the OASIS measurement tool but actually that presentation answered them. It was whether or not that added another step to the process. This is an electronically -- the tool, electronically-gathered measure.

But I think that has been addressed because OASIS is of course required. The other question that was raised in the workgroup was about the affected population, which I don't have -- I think we know that there's an improvement in the -- the measure has shown an improvement in the population. I don't have the specific numbers but we do know that this is the case.

And I think we can move on to voting unless there are questions or things that other people need to add.

CO-CHAIR GROSSBART: Any questions from the committee?

MEMBER ALMENOFF: The episode of care, this is Medicare only episode of care or is this commercial too? I just want to be clear.

MEMBER STEARNS: These are -- the Medicare population, although --

CO-CHAIR GROSSBART: Let's ask -- I mean it's used -- it's a requirement for Medicare billing and assessment. Is the tool used for non --

PARTICIPANT: Excuse me this is one of the developers from the University of Colorado. The denominator is Medicare and Medicaid patients.

CO-CHAIR GROSSBART: No commercial population.

PARTICIPANT: No. Not for Home Health Compare at any rate.

CO-CHAIR GROSSBART: Donald.

MEMBER YEALY: Okay. My question is, this looks like it was primarily developed in cardiopulmonary disease and folks with home

healthcare, and as we expand the population receiving home healthcare, that's clearly the direction things are going, will the targets actually need to change in fact? The population will likely dramatically shift in the next three to five years of who is receiving this service, therefore this particular goal, which was developed in a more narrow group, it just strikes me it may not be -- there may be a declining performance not because of anything bad happening, because you have a different population accessing that particular type of care. Am I off base about that?

CO-CHAIR GROSSBART: I'd ask the measure developer if they've got any insight on that.

PARTICIPANT: The risk adjustment process should help to compensate for that.

DR. WINKLER: To our folks from Colorado, are you on speaker phone because you are cutting out a lot of it, so we are hearing

about every third word, so if you could go to a landline it would be easier.

PARTICIPANT: Is this better?

DR. WINKLER: I think so.

PARTICIPANT: Okay. Yes, I'm sorry. What I was saying was essentially that we -- that one of the reasons for risk adjustment is to adjust for changes in the patient population that is being served, not only differences cross-sectionally among home health agencies, but also changes over time in the admitting characteristics of the population served.

CO-CHAIR GROSSBART: Dianne.

MEMBER JEWELL: But just to be clear, the measure now doesn't only focus on patients with cardiac or pulmonary diagnosis. It covers all, all eligible, other than those in the exclusion criteria, so it's already broader than just cardiopulmonary.

MEMBER ALMENOFF: I have more of a technical question. In the logistic

regression model, did I hear you say you had a robust C statistic of 0.6?

PARTICIPANT: I could not hear any of that, either of those questions.

CO-CHAIR GROSSBART: The question was, in your regression model, you had a -- your C statistic, did you say that it was a 0.6?

PARTICIPANT: About point -- Deborah, I believe you said 0.7.

MS. DEITZ: Yes, it's 0.703.

PARTICIPANT: Point seven, right.

MEMBER ALMENOFF: So that's kind of similar to a Medicare member. I mean, if you look at the recent JAMA article they actually talk about C statistics being over 0.85 or even 0.9, so your 0.7 is kind of common, what you see in the administrative data model.

But somebody used the word robust. Can I just say I didn't think that was probably a good word to use. Maybe marginal.

PARTICIPANT: I'm sorry, I am still having a lot of trouble hearing.

CO-CHAIR GROSSBART: The committee commented that they felt that the C statistic of 0.7 was maybe not robust but moderate. So with that, what I'd like to do is move us forward into the discussion and voting on each of the elements. So Christine, impact would be the first aspect of this.

MEMBER STEMPLE: Did we hear the historical trend in performance? I heard that some are improving but I didn't hear specific performance about this measure.

CO-CHAIR GROSSBART: Performance gap you mean?

MEMBER STEMPLE: Yes.

CO-CHAIR GROSSBART: We'll get to that --

MEMBER STEMPLE: Okay.

CO-CHAIR GROSSBART: -- as we go through the voting.

MEMBER STEARNS: We'll discuss it.

We'll discuss it when we get there.

CO-CHAIR GROSSBART: So impact.

MEMBER STEARNS: Oh, impact. This is reported to have a significant impact on patients with 70 percent reporting that it -- dyspnea interferes with their activity.

CO-CHAIR GROSSBART: Do them all three at once or one at a time? Okay. Okay then. So let's move on to performance gap.

MEMBER STEARNS: Okay, well you raised an interesting question about the -- the measure sponsor might be able to give us more information. There's no indication of how much improvement, just that we have an indication that there is an improvement over time.

MEMBER STEMPLE: And I guess that's my problem. If someone is getting home health, they are going through an episode of home healthcare assumingly, since it's an episode, their home healthcare care has ended, so clinically they have improved and dyspnea

is independent --

MEMBER STEARNS: Oh, let me be clear. There's an improvement in the rates so that over time, that the rate has shown an improvement. So not that the -- each individual in their episode of care has shown an improvement, but rather that the rate --

I mean I see --

MEMBER STEMPLE: Oh well, I'm struggling with this. If I'm getting home healthcare, recently discharged from the hospital, I'm improving over my 30 days of care at home, so I'm just struggling how this is -- dyspnea is a standalone, independent major cost to everybody getting home healthcare, shows much of anything that we are -- because I'm not hearing its focus, they are not -- you know, if I'm at home from a hospital I'm getting PT or therapy or whatever, I'm assuming I'm going to generally improve so I'm struggling how this independently shows some improvement in home

healthcare activity vis a vis their getting the IV antibiotics, they are getting physical -- you know, I'm just struggling, across the whole gamut of home healthcare, why are we picking this one as sort of an independent indicator that home healthcare is good?

I just struggle with the validity of the measure. Sorry, but anyway.

CO-CHAIR GROSSBART: Well, the gap in the details, the average improvement is 58 percent, so obviously 42 percent of the home care patients are not getting improvement during their home care episode or encounter -- episode.

MEMBER STEMPLE: Right. But again we don't know what the background incidence of COPD or -- we don't have the background on any -- it's just so generic, I just struggle. Anyway.

MEMBER EDELMAN: But the question is not individual improvement, but improvement of providers. I mean the point is to get a

provider to do better, I mean the point about some people will never be less short of breath, some people will always be less short of breath, so looking at individual improvement is meaningless. The issue is, does this make certain providers improve their performance, and we don't know that yet.

CO-CHAIR GROSSBART: And again, the measure developer, if you've heard that question, could you please respond? Has there been an improvement over time? Are providers changing their behavior?

(Alarm sounds)

CO-CHAIR GROSSBART: And that was our 15-minute timing. Abt, are you still on the line?

MS. DEITZ: Yes I believe that the last time this was discussed with NQF we did present some information about the fact that we have seen improvement over time for this measure, but I do not have the statistic in front of me right now.

CO-CHAIR GROSSBART: And again the documentation provided a fairly significant performance gap. Christine, finally, the evidence?

MEMBER STEARNS: Evidence. There's no specific evidence or guidelines. There's only guidelines for the treatment of pulmonary rehab, although I thought in the opening statement, there seemed to be a reference to sort of guidelines and studies that didn't seem to be in the information that I have.

DR. WINKLER: Just, this is an outcome measure, so what we are doing is looking for you know, processes of care that are likely to impact that outcome, rather than the detailed quality, quantity and consistency that we would look at for a process measure.

CO-CHAIR GROSSBART: So, point of information Reva. Do we actually have to vote on the evidence?

DR. WINKLER: Yes.

CO-CHAIR GROSSBART: Okay. Let's go through to the voting then. Well first we'll have questions.

MEMBER LEVY: Yes, I'm just trying to understand this. Maybe that's just as well, I think, for this measure. I -- so there's no evidence, according to what's submitted, there's no evidence of any process measures associated with this -- with dyspnea. Is that correct? Is that correct? Because that's what it looks like here. Which accounts for exactly what Norman was saying.

It's self-reported dyspnea, for which there are no process metrics. Okay. Just wanted to make sure I was understanding it.

MEMBER STEARNS: That's what I read.

DR. BURSTIN: This is quite typical for patient-reported outcomes. You may not necessarily have anything along those lines but this is, and so that's why there's a

bit of a pass on evidence, just a rationale for the outcome is really all that's required.

So they give a fair amount of evidence of the importance of it to patients, the importance of it to nursing, etcetera.

MEMBER STEARNS: Well, and on that point, there is a reference to -- well, the guidelines aren't specific to the treatment of shortness of breath. There are other guidelines that they used that were broader and I'd refer you to the page but -- forgive me.

So it is not that they -- so that there is that reference.

CO-CHAIR GROSSBART: Again, Dianne.

MEMBER JEWELL: So, when I first read this measure I was circling around all these same questions, and the principal reason was because the population wasn't specified enough, because there is some evidence related to the COPD population.

I hear your point about, you know, we want the providers to do better, but the problem is this is an outcome measure, and these people, you know, there's some proportion in this outcome that could just be getting better regardless of what the provider's doing, which I think is the point.

So, so the only way I can see to fix that problem is really to be more specific about which population's in-home healthcare we are talking about, not all of them, not every possible patient, because otherwise it's really more measuring did the provider document an OASIS, not did they document a meaningful outcome.

MEMBER STEARNS: Well, and that gets back to the point of it would be helpful if we had some additional information about how the measure has been doing over time, but there is a reference in here that suggests that, and it says that there has been improvement in measure over time, suggesting

that agencies are improving care for this outcome.

So, to your point, we don't have specific statistics unfortunately, and I think that we did discuss that in the workgroup, which would make it easier for us to sort of reach a conclusion on this, if we could sort of see some data.

CO-CHAIR GROSSBART: If I understand the OASIS database, and we can ask the measure developer, you're basically working with your home care patient and you are assessing them through the OASIS tool, which then helps you develop the course of therapy and care.

So the -- you are collecting this information for care delivery not for reporting, and so it may be conducive to taking steps, known practices to improve care. Again, the submission was a little thin on the evidence and only, I believe, cited two studies. Can Abt comment on that?

MS. DEITZ: I'm sorry are you looking for developer input?

CO-CHAIR GROSSBART: Yes.

MS. DEITZ: I'm sorry, it's a little hard to hear. There is basically -- I think, are you asking us why there are only two studies cited?

CO-CHAIR GROSSBART: No, we are asking what is the evidence that having this measure is going to impact processes performed by providers.

MS. DEITZ: That is going to? I mean --

CO-CHAIR GROSSBART: Well, I mean, so, I mean, how is a provider going to use -- I mean what's the evidence that there are things the providers can do to change these numbers?

MS. DEITZ: Well, as I mentioned, there are best practice packages that have been put together by the quality improvement organizations, and agencies have adopted these

as guidelines to improve their practice, so that -- they select a measure, an outcome measure like improvement in dyspnea, and know if their measure seems below the national benchmark, and choose it as an area that they want to improve on, and then they use the best practice package to improve their practices and note whether or not their patients are improving on this outcome.

As context I just want to mention that CMS reports -- publicly reports, NQF endorsed measures on a variety of outcomes, such as improvement in ability to -- in speech and language, improvement in level of pain, improvement in ambulation.

This is the measure that addresses dyspnea.

CO-CHAIR GROSSBART: Dianne.

MEMBER JEWELL: Can you hear me? I'm talking to the measure developer. Can you hear me?

MS. DEITZ: Yes.

MEMBER JEWELL: Thanks.

MS. DEITZ: So can you help us understand why the population for dyspnea in this measure is all home healthcare patients as opposed to say, patients with COPD and heart failure, for whom the clinical indication of dyspnea is much more specific?

MS. DEITZ: I believe that one way to think about this is that the population of home health patients frequently has dyspnea as part of their -- the experience of having been chronically ill for a variety of conditions, and having been bed-ridden and deconditioned, and so -- and then the dyspnea then, you know, keeps them from engaging in activities that could improve their -- the -- and decrease their shortness of breath.

So the idea is that it's not just patients with specific conditions that need to -- attention to their dyspnea.

CO-CHAIR GROSSBART: I have a question. Would a new mother who is on

Medicaid get a home care visit and have this data element collected on them, through OASIS?

MS. DEITZ: Well, if it's a maternity patient, the answer is no, but the -- around pre- and post-maternity care, are not included in the OASIS.

CO-CHAIR GROSSBART: Well with that, I think we should move on to the voting, keep this going. So importance of the measure to report and impact. One if it's high, two if it's moderate, three if it's low, and my eyes aren't good enough to see that counter up there but it still looks like it's in the teens.

(Pause for voting)

CO-CHAIR GROSSBART: How many votes are we at? There we go. And so the vote is seven moderate, seven low, and six insufficient evidence. We're done.

Okay, so with that the NQF has moved to -- committee has moved to recommend non-endorsement. And Kevin, you're back on.

CO-CHAIR WEISS: Okay, we're back on to the AMA PCPI measure on empiric antibiotic use for --

DR. CANTRILL: I'm Steve Cantrill. I'm an emergency physician from Denver and was involved in the original multi-disciplinary group that I believe was in 2006, helped develop these measures, and have been asked to at least provide the introduction, although I have a lot of support here from the folks that actually know all the data.

We were talking and I would ask a favor of the Chair, since we are running late, could we possibly do all four of these in a row, because I have to catch a flight back to Denver from Dulles this evening?

CO-CHAIR WEISS: We'll check on that while you are doing this one. We just have to make sure everyone else isn't queued up in a weird way.

DR. CANTRILL: Thank you. Actually I'm going to give the introduction

for all four.

CO-CHAIR WEISS: I think the answer is yes, because I think we're okay with the other ones. So go -- let's plan on it.

DR. CANTRILL: Good thank you. In terms of introduction to all four of the measures, we are talking about empiric antibiotic therapy in community-acquired pneumonia, CAP, for patients that present to the emergency department, both those that are discharged and those that are admitted.

We're also talking about vital signs that are recorded and reviewed for patients presenting with CAP, assessment of oxygen saturation both recorded and reviewed for patients with CAP, and mental status evaluation for patients with CAP.

These were, as I said, originally developed and approved by PCPI in 2006. They were endorsed by NQF in 2007. We are here for endorsement maintenance for the first three. The fourth measure, the mental status

evaluation in fact was -- the endorsement was removed in 2010 because they felt there was not a performance gap, although the latest data that we have, unfortunately it's from 2008, from PQRS, these measures were all part of PQRS from 2007 through 2012.

Those data demonstrate a gap of 23 percent in terms of empiric antibiotic therapy, 22 percent for vital signs, 20 percent for oxygen saturation assessment and 19 percent for mental status evaluation.

It's because of the 19 percent, we feel that is a significant gap, that's why that measure is being resubmitted, as you can tell by the number, 1895.

These, all four of these measures have been tested for reliability and validity, and as I mentioned, have been part of PQRS from 2007 through 2012.

If I could just address a couple of the items that were brought up by the steering committee, in terms of Measure 0096,

the empiric antibiotic therapy, there was some concern about having treatment for atypicals as well as bacterial pneumonia, and the Cochrane study, 2010 Cochrane study was mentioned. My concern about that study, I understand there were three papers that really dealt with this issue, two of which I believe were in Europe, where atypicals are not as much of an issue, and all had relatively small sample sizes. So I think that may be something to watch but I think that we are consistent with the IDSA/ATS guidelines and so we feel comfortable with that.

In terms of Measure 0233, assessment of oxygen saturation, the question was should a timeframe be specified. Every emergency department I've ever been in, O2 sats are the fifth vital sign and they are the first thing that are obtained by the nurse when a patient arrives.

So you could ask for that. I don't think it's going to give you much

information. In terms of Meausre 1895, the mental status, again, a timeframe. Should a timeframe be recorded.

That is somewhat problematic since that is nominally part of the physician's physical examination in terms of mental status, and very often you are lucky if that has a time stamp of when it's recorded, let alone of when it's done.

So I think that might have a very negative impact on the feasibility of this measure. Also, the question was should a particular tool be used to determine the mental status. The question actually is not really mental status. It's confusion, and confusion is used in a couple of -- in like the CURB-65 tool in terms of determining who should be admitted versus who can go home with pneumonia, and also a proposed tool in terms of who should go to the ICU.

So it's really confusion and you don't need a tool to determine confusion when

you are examining a patient.

So that really is all I have to say in terms of the introduction.

CO-CHAIR WEISS: Okay, Dave.

MEMBER RHEW: Yes Steve, hi, Dave Rhew. Just had few questions with regards to Measure 0096. You mentioned the IDSA/ATS guidelines. I am -- there's an implicit assumption that that is, when you say empiric, appropriate use, that's the guideline that you are looking for, right?

DR. CANTRILL: Yes, it is.

MEMBER RHEW: Okay, just a recommendation. Maybe we could --

DR. CANTRILL: It's a 2007 one, is what it is. But we are still consistent with that.

MEMBER RHEW: Right. And I think we all know that, it's just it would be nice if it were actually in the document and explicitly stated so it could be easier to follow.

Additionally, the numerator/denominator, it's -- for 0096 it says essentially patients with appropriate empiric antibiotics in the numerator, and the denominator, patients aged 18 and older with a diagnosis of community-acquired bacterial pneumonia.

Does this mean that the supply is to both inpatient and outpatient?

DR. CANTRILL: It does, and that I think, you know, one of the related measures is 0147, from CMS. Now that really applies only to inpatient.

Here we are dealing with all comers and we think that is a -- that's really the measure that we need.

MEMBER RHEW: Okay, it's a total, all patients, inpatient, outpatient denominator.

DR. CANTRILL: All adults.

MEMBER RHEW: Got it.

MEMBER STEMPLE: And so is that

for all these measures? Because I'm a little confused. Are we talking all these measures are inpatient outpatient, all the ones we are discussing at this point in time, because it's not clear to me.

CO-CHAIR WEISS: Okay. That's a question. Are all these -- PCPI, are all the measures -- are all of them inpatient, outpatient?

DR. CANTRILL: Mark, do you want to address that, the way you think there might be a little confusion with that.

DR. ANTMAN: Right, at least for 0096, the intent is for it to be ambulatory or outpatient only. We are double checking on the others, but I believe the entire set is intended to be outpatient only. But again, we are double checking that.

CO-CHAIR WEISS: Okay, so we'll know that shortly.

MEMBER RHEW: So, if that's the case, maybe we could put that as in the

specifications, only ambulatory or exclude hospitalized patients.

MEMBER STEMPLE: Yes, the numerator is not clear because the numerator would seem to intend all patients regardless of site of care, so --

CO-CHAIR WEISS: Okay, so Dave, if you can give us the impact, the gap -- oh, sorry. Peter.

MEMBER ALMENOFF: Are the location of the patients, we are trying to figure out, is this patients in the emergency room, in the outpatient arena, it doesn't say anywhere, where these patients are supposed to be.

So if we are going to do oxygen saturations and vital signs, in what location?

DR. CANTRILL: Well, the intent was originally emergency departments, but I don't know what the instructions are to the hospital when they gather the data.

MEMBER ALMENOFF: It's not in here.

MEMBER YEALY: I think we are conflating like three different measures. I think two of them are ED-specific and others are not.

CO-CHAIR WEISS: So I think we have got to go through them one by one.

MEMBER ALMENOFF: It doesn't say it in any of the writeups.

MEMBER YEALY: Some of them, in the title it says emergency medicine, so by definition it means that. But not this one.

CO-CHAIR WEISS: So, again, let's focus 0096 now, if we can, so Dave, we'll go to impact, gap, and evidence.

MEMBER RHEW: Sure, so focusing on impact, first off, we all recognize pneumonia and influenza, eighth leading cause of death in the United States. Pneumonia is the number one cause of death due to infection, high cost, clearly an area where there is significant opportunity for improvement.

So that's one of the areas around

impact. Do you want me to just go through each one of them? And then we'll do the voting?

CO-CHAIR WEISS: Each one of them being -- each one of them what? Measures?

MEMBER RHEW: I'm sorry, the impact, performance gap, and evidence.

CO-CHAIR WEISS: Yes, do those as a group.

MEMBER RHEW: Okay. Performance gap, you know, this is one where we actually as a -- the group that started thinking about whether or not there was a gap, we weren't quite sure exactly when do you define the threshold for what a gap is.

So we know that the PQRS 2009 data suggest that the current use is 92 percent, in the Hospital Compare it's 94 percent. Does that mean that's a gap that's large, small? We didn't really know where to draw the line.

But we certainly acknowledge that it's over 90 percent and that may or may not

be a gap that's large enough.

Now, the evidence, I think, clearly large, observational data sets show the clear benefit, especially for those that are hospitalized, severe pneumonia patients, in particular the ICU, bacteremic, pneumococcal pneumonia ones, where combinations of macrolides on top of the beta-lactams have been shown to impact outcomes, as Steve, you noted that there was a question as to whether or not the data from the Cochrane meta-analysis of 25 RCTs is valid, and in fact I think, Katie, you sent out an email to the group which outlined several of the reasons why that study may not be applicable to this conversation, and these were data that were shown to us by Dale Bratzler and several of his colleagues.

So in addition to that, I can also add one other piece. Most of the RCTs that have been published out there are pretty much not powered to demonstrate superiority. They

are more around equivalence, therapeutic equivalence, so that would be one other thing that you could also add to the list.

So, that said, what you are left with is large data sets that show a clear association, especially for severe patients.

0096, though, is for the ambulatory and so the question then is, as the evidence starts getting weaker, that you start relying more on extrapolations and expert opinion.

So the evidence, clearly the data sets for the inpatient, especially the ICU, severe, there, ambulatory, you know, again, good, good reasons to believe, you know, that there's atypicals out there, you should cover for, but it just, we don't really have as much data on that.

So that's a quick overview of the impact, performance gap evidence. The folks that were on the review committee, any other thoughts?

CO-CHAIR WEISS: No, at least I am seeing that the group was pretty equivocal on performance gap, as measured by your earlier telephone meeting. Can you explain that a little bit to us?

MEMBER RHEW: And again, we are looking for some direction from I guess the larger group whether or not a 92 or 94 percent, you know, level is representative of a significant gap, or if that's sufficiently high enough. Don?

MEMBER YEALY: And I guess my take on it would be, is that I'd be sold that 92 percent is enough of a gap if in fact we were talking about the sickest of the cohort, but in fact this is going to be a predominantly ambulatory cohort, so 92 percent may not represent all that much of a particular gain.

That's the rub here between this and the other antibiotic criterion.

MEMBER RHEW: Yes, I think clearly the data are so strong on the inpatient for

the observational -- observational data are very strong on the inpatient, severest patients, but again, you know, 92 percent of the ambulatory, I don't know. Maybe or maybe not.

MEMBER CANTRILL: If I could just comment about. There may be some confusion because the 2008 PQRS data, I am told that the gap is 22.52 percent, which is obviously different than the 9 percent, and I don't know where this confusion -- can we elucidate this?

DR. ANTMAN: I apologize if there is some lack of clarity in our submission forms. These measures are intended for use in the ambulatory setting. Again, that may include the ED, but these do not include inpatient.

CO-CHAIR WEISS: So, it says in here that this is reported as part of Hospital Compare so I'm confused again, if it's ambulatory, why is it part of Hospital Compare's reporting?

MEMBER STEMPLE: Or at least, I'm looking on page six.

MEMBER RHEW: Yes, I think at that point when we were -- we weren't sure at that -- we just learned that it's now ambulatory. We weren't sure if it was a combination so we had included the Hospital Compare, but at the bottom, you also will see, during the call I'm not sure who mentioned that they had looked at the 2009 PQRI data and they said it was 92 percent. So I don't know where those came from or if those are actually correct.

CO-CHAIR WEISS: Is there a way to quickly look at the PQRI data while we are --

DR. WINKLER: There it is. It's on page three of the submission form. This is the 2008 data, 10th percentile is 33 percent, 90th percentile is 100 percent, 50th percentile is 90.9 percent.

CO-CHAIR WEISS: So we are seeing a median of 90 -- 91 percent in an ambulatory environment, as our number, with some

variability, and maybe it's not the absolute. Maybe it's the variability that is of concern here.

Okay, any other issues related to gap or evidence? I did see on the workgroup as well, there was equivalence on the evidence as well, and I just want to make sure that we all understand what that equivalence was, because it was not uniformly high. It was medium and lows in there.

MEMBER RHEW: And I think at the time we were still trying to reconcile what to do with the Cochrane review and the data and I think since then we have obtained some really good feedback from the developers, and CMS, and Dale and others, so I think given those caveats, you know, and if we are going to accept those as reasons that we wouldn't include the Cochrane, then I think what you are left with is still observational data and there is a need to do a large, randomized, multi-center control trial, but in the absence

of that, the data would suggest that it's beneficial at least for severe patients.

CO-CHAIR WEISS: Thanks Dave. Let's go into a vote, if there are no other questions. So impact, one high, two moderate, three low and four is insufficient evidence for impact.

(Pause for voting)

CO-CHAIR WEISS: Press those buttons good, or well. Is it good or well? Oh, just press them hard. Push, push well. Okay. Push again. There it is, okay good.

So we've got 11 high, 8 moderate, 1 low, no insufficient evidence. Good. Next. Let's go to gap. High, moderate, low, one, two, three and four is insufficient.

(Pause for voting)

CO-CHAIR WEISS: Okay. Almost there. There we are. Two high, 12 moderate, 4 low and 2 insufficient. So kind of milquetoast about this, but okay.

Okay. Sorry. 1c, evidence. This

is just a yes, no. Do we feel there's enough evidence that supports this, or is there insufficient evidence?

(Pause for voting)

CO-CHAIR WEISS: There we go. Yes, 15, no 1, and 4 believing there is insufficient evidence. So let's go forward and go to reliability and validity.

Dave.

MEMBER RHEW: Yes, clearly this has been tracked already so we know that this can be tracked, and it can be tracked reliably. So I would say that our thoughts were yes, it's reliable.

CO-CHAIR WEISS: And validity? We spoke a lot to that issue already but are there specific issues you'd like to raise up?

MEMBER RHEW: I think the one thing about that was, and again, this applies not only to this but to the next measure, a large number of these metrics are based on time that you see the patient, but really when

we measure it, it's based on the final diagnosis.

And there's a disconnect there. I mean when a patient comes in, you don't know they necessarily have community-acquired pneumonia. They come in with shortness of breath and then later on you find out that they had pneumonia.

And then the way that you evaluate it though, is all based on, well, they had pneumonia, at the final discharge diagnosis. So, there is a disconnect there that we struggle with.

(Alarm sounds)

MEMBER RHEW: And I think we just wanted to acknowledge that.

CO-CHAIR WEISS: And acknowledging it in what -- as a neutral force or as something to be concerned about or --

MEMBER RHEW: It does create some challenges. We can't quantify whether or not it, you know, makes it invalid or you know,

but it certainly makes it harder for us to determine whether or not we have captured all the patients with pneumonia, because it's all entirely dependent on the clinician's ability to properly diagnose what at the time that they are seen.

So it creates some variability and some questions as to whether or not this -- that may influence the results.

CO-CHAIR WEISS: Okay. Peter and then anyone else also in the workgroup that would like to comment here.

MEMBER ALMENOFF: Well, we are all over here all kind of a little challenged about this measure. We can't tell if this is an inpatient or an outpatient measure. One time we are hearing the word discharged and the next minute we are hearing ambulatory. We -- it doesn't say anything in here what this is.

So I think we need to establish is this an inpatient measure or an outpatient

measure? Of all these, I mean I --

CO-CHAIR WEISS: Are the specifications clear enough for us?

MEMBER ALMENOFF: It's got to say it somewhere and it's got to be written down because they are completely different things.

CO-CHAIR WEISS: Mark, do you want to help us here?

DR. ANTMAN: If I may, if you look at the specifications, and I'm not sure --

MEMBER ALMENOFF: We don't have all that. We only -- I only have this sheet in front of me and it doesn't say anything about inpatient or outpatient on any of these.

So I just need to know.

CO-CHAIR WEISS: Let's pull up the specs and see. When we are doing that, any other thoughts or comments?

MEMBER PELLICONE: If it's really outpatient care, the implication is that every one of these patients gets a chest x-ray, and that's -- if you're talking just -- if you're

talking emergency room that's one thing, if you're talking an office or an outpatient clinic, that's a huge burden, if you go by the classic teaching that you need a radiographic infiltrate to make the diagnosis of pneumonia.

CO-CHAIR WEISS: Okay, is this the actual specs?

MS. WEBER: And the measures were actually on the thumb drive as well.

CO-CHAIR WEISS: Electronically or by the -- Okay so can someone help us here walk through this? This would be great. The question is, is this inpatient or outpatient or is this --

DR. WINKLER: The numerator says patients with appropriate empiric antibiotic prescribed, numerator time, one for each episode of community-acquired pneumonia during the measurement period.

In the details it just says this measure should be reported once for each occurrence of pneumonia during the reporting

period.

Definitions, it doesn't really say anything about setting. EHR, nothing. So in the denominator, all patients 18 years and older with a diagnosis of community-acquired pneumonia, time window, each episode of CAP, really does not say anything.

The denominator details, it lists the EHR, I mean the claims or administrative codes, patients aged 18 years and older. It has the ICD-9 diagnosis, some CPT II codes and then an asterisk, it says clinicians using the critical care code 99291 must indicate the emergency department place of service on the Medicare Part B form.

But okay.

DR. ANTMAN: Reva, may I -- so again, any confusion about the setting of care is certainly not intended on our part. And if it would be helpful if we added language to clarify that the intent is for these to be ambulatory only, language to the denominator,

we can certainly do so.

I think our feeling was that it would be fairly clear in that the CPT codes that you just referenced for the denominator, Reva, those are all ambulatory visit codes, with the exception of those codes that are applicable to the ED.

CO-CHAIR WEISS: So, I need a process question answered by staff, and that is are we allowed to accept in this meeting an amendment by a developer as part of our process, or does that have to happen outside? I don't --

DR. WINKLER: I think if it's a clarification, that clearly we can -- that they need to put some additional language to make it clear to respond to some of these uncertainties and questions.

It isn't a change in the measure. It's more a matter of just making it clear for everybody's common understanding.

CO-CHAIR WEISS: Is that what we

are doing here? And Mitchell.

MEMBER LEVY: Can I ask the -- why is this limited to the ambulatory setting? Is there some reason that we think that it's more important for people to get appropriate antibiotics in an ambulatory setting, including the ED, as opposed to inpatient?

DR. CANTRILL: Well, the next measure is focused on inpatient.

MEMBER LEVY: Oh, the next one --

DR. CANTRILL: In ED. Yes.

MEMBER LEVY: I thought both were -- okay. All right.

MEMBER ALMENOFF: And to me it doesn't matter. I just want to know which setting, and if nobody can even answer this question --

CO-CHAIR WEISS: So let's go forward now --

MEMBER ALMENOFF: It worries me that nobody can answer this question.

CO-CHAIR WEISS: So Peter let's go

forward with the presumption as we have been told by the developers that this is an -- that the word outpatient, or at least non-inpatient is what we are seeing here. Right? Is that correct?

DR. ANTMAN: Correct.

CO-CHAIR WEISS: Okay, so let's continue forward with this last little bit of discussion on reliability and validity as a non-inpatient.

MEMBER LEVY: But it would include emergency department.

CO-CHAIR WEISS: Non-inpatient, which would include -- okay.

So with that in mind, Dave, we are back to you again, on validity.

MEMBER RHEW: Yes, I mean again, some of the things that we have already mentioned, I think the only other thing is the IDSA/ATS guidelines calling that out explicitly so we know exactly that, you know, because nowhere do they mention any

antibiotics. It's just assumed that it's the current 2007 IDSA/ATS, but beyond that, I mean, we certainly know that this has been captured, it is easy to capture through the EHR. So as long as it's, you know, given the caveat that it is still retrospective.

CO-CHAIR WEISS: Okay. Good. So any questions or comments from the workgroup on reliability or validity?

(No response)

CO-CHAIR WEISS: Any from the table at large?

(No response)

CO-CHAIR WEISS: Then let's go to vote. So we are looking at reliability, one, two and three, high, moderate, low, or insufficient.

(Pause for voting)

CO-CHAIR WEISS: And we are up to 16, or 18, 20. There we go. Okay. So 7 high, 11 moderate, 1 low, 1 insufficient. Let's go on to validity.

One, two, three and insufficient again, so high, moderate, low, insufficient.

(Pause for voting)

CO-CHAIR WEISS: Okay everyone press again please. Squeeze that last one out. Okay is everyone pressing again? Let's try, everyone again, one more time.

Pretty soon we are going to have to be focusing on Jessica here. Okay smile Jessica, everyone focus to Jessica here. I'm not sure if this is magical or not. Who knows?

No. Someone is not doing theirs and we are going to find out who that 20th is. Well, it should go anyway at the end of the time, right? To okay. So we'll do 19 and this one looks like -- okay. So 4 high, 13 moderate, 1 low and 1 insufficient. Does that come up to 19? Yes. Okay.

Oh, actually are we missing some between Mitchell and Michael? Was there someone sitting there?

No? Everyone's here. Okay good. Good, let's go on to usability and feasibility. Let's try and make these crisp, clear and succinct now that we know that this is a non-inpatient measure.

Correct. So --

MEMBER RHEW: Again, those were from the workgroup and we didn't have a clarification at the time that it was ambulatory. We were just looking at it and it looked like it was all comers. So that's the only reason it's there.

CO-CHAIR WEISS: Okay. Does that help you? So usability. Can this be meaningful and understandable? (Laughs)

Sorry. That was not meant as an editorial laugh. Meaningful, understandable and useful for private, for public reporting and accountability and the same thing for quality improvement.

So you're the workgroups --

MEMBER RHEW: Again, assuming --

we are assuming that all those things that we talked about have been incorporated into this.

CO-CHAIR WEISS: That's a lot of assumptions but we are told that those are just clarifications.

MEMBER RHEW: Okay.

CO-CHAIR WEISS: And what was the answer to that?

MEMBER RHEW: I mean we did feel that -- yes. Assuming all those things were in place, yes you would be.

CO-CHAIR WEISS: Okay, good. And then from the workgroup, any comments on usability beyond what we have heard from Dave? And then from the table at large?

(No response)

CO-CHAIR WEISS: Let's vote on usability. One high, two moderate, three low, four insufficient information.

(Pause for voting)

CO-CHAIR WEISS: Okay, let's vote everybody. Please vote. We're up there.

Okay. Are people's thumbs getting tired? Is that what's happening here? We are going to do some thumb exercises in a few minutes. Let's do it. Up, down, up, down.

Okay. Everybody let's shake your wrist, make -- get real comfortable and let's try it again.

There we go. Somehow we got it. Someone got it. Okay. So it was a split vote of eight high, eight moderate, two low and two insufficient. I was going to say four, but I didn't.

Feasibility. Here we go Dave.

MEMBER RHEW: Again, since this is currently being collected through the EHR and other mechanisms, we felt it was definitely feasible and again, assuming all those other caveats, it will be applied.

CO-CHAIR WEISS: Great. Workgroup, any addition to anything Dave said? Table at large? Anything?

(No response)

CO-CHAIR WEISS: Let's vote for feasibility. One high, two moderate, three low, four insufficient.

(Pause for voting)

CO-CHAIR WEISS: We're getting there. Almost. Okay. That's it. There we go, 15 high, 4 moderate, no low and 1 insufficient information.

Let's go to the final vote for the measure, summative. Yes, no. Should we move this forward towards endorsement? Please vote one yes, two no. All the clarifications are made. We can't call them modifications, because that would be wrong.

(Pause for voting)

CO-CHAIR WEISS: So 18 yes, 2 no. Let's continue on now with the next measure.

DR. WINKLER: Now the question is, do we have the folks from CMS on the line for Measure 0147 0148?

We have had a request to move the other PCPI measures ahead. Is that a real

problem for you all?

DR. BRATZLER: This is Dale Bratzler. I can wait a while

DR. WINKLER: Thanks Dale. I appreciate it very much.

CO-CHAIR WEISS: That sounds great. So we are going to go to Measure 0233. I am going to do something here that I think might be helpful to the group as well. I am going to suggest we do a standup break at our place.

So this does not mean we leave the room, unless you absolutely have to, but really it's just to stretch your legs and sit down again.

We'll do a real break in a few more minutes but let's just all just take a standup and please all just stand for just a second, we'll all feel better for it. Norm, a little standup here, Christine, just shake it around a little bit.

This is not meant to be a chance

for everyone to go skedaddle. Okay? Okay. Sit as you feel comfortable to sit. Dave, we are ready to go, 0233. So we have our measure developers who have spoken to -- it's my fault. I got you all kind of moving in here, got that blood moving.

MEMBER RHEW: Okay, Dave. Were we going to give Dale an opportunity? Oh, we got him at 0147? Or which one are we --

CO-CHAIR WEISS: We're going to 0233. We are going to be doing -- oh sorry. John. Okay. So we are on 0233. It's John. I apologize.

We've already had the measure developer speak to us, so we are going to go right into impact , gap and evidence.

MEMBER PELLICONE: Well, I think we have heard about the impact with regard to community-acquired pneumonia. There's also significant evidence that the degree of hemoglobin O2 saturation is of great significance with regard to morbidity,

mortality and the ultimate destination of the patient, if they were going to be hospitalized.

I did have a question for the developer about the report of the hemoglobin O2 saturation. Is it understood that the O2 saturation is always to be reported with the FiO2? Because there are great implications here with regard to --

(Alarm sounds)

MEMBER PELLICONE: PQ mismatch and AA gradient with the report.

DR. CANTRILL: Clinical gradient is you have to know the FiO2 to make any sense out of the 02 sat.

MEMBER PELLICONE: Yes, I just unfortunately see too many instances in which it's not --

DR. CANTRILL: Well, what should be and what happens are two different things, as you know. We try to always have the FiO2 specified.

MEMBER PELLICONE: Okay thank you.

CO-CHAIR WEISS: Is that in the specification is the question? Is that in the specification? Can we ask the developers to take a look and see how well that's specified in the specification, and let's continue while we look at that.

MEMBER PELLICONE: With regard to the performance gap, there was the report about the PQRS study that there is about a 20 percent performance gap there, so there's -- I think that is of significance.

And with regard to the evidence, there's -- there's level two and level three reports based on the ATS/IDSA reports as well as the 2001 ATS guideline, in which this hemoglobin O2 saturation value has been studied. So that's it.

CO-CHAIR WEISS: Rest of the workgroup would like to comment on anything that John has said or anything else that you feel happened, Don?

MEMBER YEALY: I am even a little less bothered about the evidentiary gap, because not only is it incorporating virtually every risk stratifying score, in the most prominent scoring system, the PSI, it actually gets counted twice, I mean, it's incorporated into the numeric score, but if you're hypoxemic it doesn't matter what class you are, you can't -- you have to be non-hypoxemic, one through three, to really be low.

So you couldn't have anything more basic than this. So the evidence in my view is actually overwhelming, and what's frightening is, I heard people talk about 100 percent pediatric instructions, which seem to me to be a lot to do, and we can't get a fingertip probe on somebody with -- when the target end organ is the lung.

I'm always amazed that this remains an opportunity.

MEMBER RHEW: My take is that we

are missing a critical piece, which is the timing, i.e. you have to specify that it needs to be done within a certain time period, whether it's one hour, three hours, or whatever --

MEMBER YEALY: And I think he's just saying it's in the emergency department, and even at that, it's hard to -- and I realize how that sounds, that sounds incredibly average, length of stay for an ED visit is about four hours give or take, that's a pretty wide swath to still find 20 percent failure.

MEMBER RHEW: Yes, but I would say in terms of the ability to impact the outcomes, if you can do it within one hour, versus -- and I know in some EDs you can wait almost up to 24 hours, or you know, you'll be sitting there all day.

I mean there's a huge difference in terms of your ability to impact care, so I would strongly suggest that we add a timeframe

to this.

DR. CANTRILL: If I could comment on it. David, I mentioned, you know, it really, at least in emergency medicine, it's becoming a fifth vital sign. And what's the first thing that happens when a patient comes into the ED? They get a set of vital signs.

So we could specify that, but I think it would complicate unnecessarily the measure, because vital signs are always taken when the patient appears, so you -- before I go into see the patient I've got, you know, blood pressure, pulse, respiration and O2 sat.

CO-CHAIR WEISS: So, Mitchell.

MEMBER LEVY: But if that were the case then we don't need a performance metric. So we are saying two separate things here, I mean the 10th, the 25th percentile is 71.43 percent, so you can't have it both ways.

DR. CANTRILL: No, I understand. I understand. But if it's going to be done, it's going to be done up front.

CO-CHAIR WEISS: So, let me just ask, is there a perception that this is not being done at a high rate, or is a documentation problem why we are seeing the low rates? What is -- what have we learned about this measure?

MEMBER YEALY: My answer would be yes. To both of those actually. And the bottom line is you won't -- without solving both, you won't actually be able to address this.

CO-CHAIR WEISS: Has there been any work done by the developers to understand this gap as to what is going on? Is it a documentation gap or is it a care gap?

MEMBER YEALY: All I can tell you is inside the ED community-acquired pneumonia trial in RCT, almost 10 years old now, that one of the quality metrics was please measure oxygen saturation and we'll pay you by patient for it. It still didn't hit 100 percent. So --

CO-CHAIR WEISS: But where was it though?

MEMBER YEALY: In the emergency department.

CO-CHAIR WEISS: No no, I'm sorry, where was the metric? It didn't hit 100 percent but was it at 80?

MEMBER YEALY: Ninety-plus.

CO-CHAIR WEISS: Ninety-plus. Okay. And that was an additional incentive to do it?

MEMBER YEALY: You were inside of an RCT. You would think if any place -- and the whole RCT was about approving process care -- if any place you are going to hit 100 percent, that would be it. We didn't.

CO-CHAIR WEISS: So any more questions on impact? On -- okay, let's measure. Impact.

High, moderate, low, insufficient is four.

(Pause for voting)

CO-CHAIR WEISS: I'm just thinking

how they are doing all these good indices when they don't have the saturation to do them. Fifteen high, four moderate, one low, and no insufficient.

I was just going to say four insufficient. Next. Performance gap, we have talked about it, is there any more discussion you want to have on performance gap? Anyone want to raise anything?

(No response)

CO-CHAIR WEISS: No. Okay. Let's go to -- yes.

MS. CHAVARRIA: And this is just going back to Dr. Rhew, what you had mentioned before, on the call we had -- we had mentioned the 2009 data and we had mentioned that there was no variability so we didn't know what -- in the 2009 data what was provided to us, we didn't know what the variability was.

But I do have it for this particular measure, for 2009, we only have the mean measured reporting rate, and out of

203,500 eligible professionals, there was still a reporting rate mean of 86 percent, and that's for 2009.

MEMBER EDELMAN: I'm sorry, did you say professionals, or EDs?

MS. CHAVARRIA: It was for the eligible professionals.

MEMBER EDELMAN: Does that go -- does that include the doctor's office?

MS. CHAVARRIA: Yes.

MEMBER EDELMAN: But I thought we were talking about an ED criterion. I'm confused. Are we talking about the doctor's office or the emergency department?

DR. ANTMAN: So this measure is reported at the level of the individual clinician, so if it is being reported from the ED, it is being reported by an individual physician.

MEMBER EDELMAN: I know, but what if it's being reported from the doctor's office? Is that included in the statistic you

just gave us?

DR. ANTMAN: Yes. That would cover --

MEMBER EDELMAN: So is it or is it not relevant to what we are considering?

DR. ANTMAN: Forgive me Dr. Edelman, is what relevant?

MEMBER EDELMAN: The statistic we just heard includes a doctor in his office making a diagnosis of pneumonia.

DR. ANTMAN: Yes.

MEMBER EDELMAN: And whether or not he recorded oxygen saturation. It's my understanding the metric we are considering applies only to emergency departments.

DR. ANTMAN: That's not correct. The --

MEMBER EDELMAN: The metric we are considering applies to physicians in their offices?

DR. ANTMAN: Right, to use Dr. Weiss's language, the non-inpatient setting,

which can be any ambulatory setting, including the ED.

MEMBER EDELMAN: Okay. So I am confused because I thought I heard something else from Dr. Weiss actually. So can we decide what we are talking about?

CO-CHAIR WEISS: So is this measure also a -- so we are on 0233, it's titled emergency medicine, assessment of O2 sat for CAP essentially.

So is this an emergency medicine --

MS. CHAVARRIA: So we did -- we had -- the updated one that we had, that probably NQF staff put up on the website, it had the removal of the emergency medicine piece on it, because -- so now it's just assessment of oxygen saturation because in the first once since it had been required of the emergency set, that was included in the title.

CO-CHAIR WEISS: So this is now for non-inpatient.

MS. CHAVARRIA: Yes. And when we submitted that update, it did make it into the updated form, but perhaps you were working off-of, perhaps the --

CO-CHAIR WEISS: Okay, and then that would mean that this is a -- 86 percent includes all emergency room use of this and outpatient, which means you have probably got apples and orange things going on, which is probably close to 100 percent in emergency rooms.

And actually, that was --

MEMBER EDELMAN: But they're probably quite low. When you come into your doctor's office, and you have a little fever and you're coughing a little bit, and he hears a little junk in your chest, and he says you've got pneumonia and he gives you a Z-Pak, the odds are he's not going to do an oxygen saturation if you look well.

And the question is do we want that?

CO-CHAIR WEISS: Yes, and so is the intent that every outpatient, every physician in their office should be doing an O2 sat before a diagnosis of -- and treatment of a patient with CAP, is the question I think we are moving ourselves to. Is that --

MEMBER EDELMAN: Yes, no, the question is entirely different now, and I think it deserves a little consideration.

MEMBER PELLICONE: I also have another issue, is the patient arrives in the emergency room and is evaluated, diagnosed with pneumonia and sent home, they are included in the measure, but if they get admitted they are excluded from the measure, so we are splitting the ED visits. Is that correct?

CO-CHAIR GROSSBART: But they would be included in the current O2 assessment for inpatients.

MEMBER PELLICONE: Which is not, which is not -- lost our endorsement. Which

is not part of this because it's been at 100 percent for about three years.

CO-CHAIR WEISS: But you have to come back to what we are saying here though is you'd want to have every physician practicing pneumonia treatment --

MEMBER EDELMAN: Basically we are asking every physician who ever makes a diagnosis of pneumonia to do oximetry.

CO-CHAIR WEISS: Yes. Is that what --

MEMBER EDELMAN: Is that the standard -- is that the standard you are proposing?

DR. CANTRILL: That is the standard we are proposing.

MEMBER EDELMAN: Okay thank you.

DR. CANTRILL: The location should have no bearing on care, in terms of -- and it has become a fifth vital sign.

MEMBER EDELMAN: Yes, I understand that. But you have to understand physician

behavior in private practice, and you know, a cough and a little fever will frequently generate a diagnosis of pneumonia.

Now, you know, I'm not sure requiring oximetry is appropriate.

CO-CHAIR WEISS: But that goes back to the evidence, and we approved the evidence. We are now on the -- oh actually we haven't approved the evidence. We only did -- you're absolutely right, we only did impact.

Okay, so we see a performance gap, the performance gap we now understand better, which has to do with a -- it's a concatenated performance gap of emergency department and outpatient, and it's running at 86 percent, which we would suspect, no evidence, I mean no real data show that it's going to be very high in the emergency room, which means it's going to be very low in the community, which is not inconsistent with our gestalt, with those of who have a sense of this.

So let's now vote on performance

measurement gap, is there a gap? High, moderate, low and insufficient evidence.

(Pause for voting)

CO-CHAIR WEISS: This one has gotten a little complicated. I didn't anticipate it. But 19 of us have voted. Maybe 20 of us have voted and it just hasn't -- we'll vote again. Don't change your minds. There we go. Done.

Stop voting everybody here please. Thirteen say high, four say moderate, three say low, and three -- zero say low and three say insufficient. I wonder if it's just me or if anyone else who was up here would do the same thing. We'll see. We'll see. It's probably me.

Okay. Evidence. Now we are to the evidence, and so we have been discussing a lot of this evidence question about what this means in the non-emergency setting, and what's the evidence about this.

And what we are hearing from our

measure developers as I understand it, is that the evidence would say, regardless of where you're diagnosing pneumonia, you need to get this O2 sat.

Now one of the questions we were asked was do we also get an FiO2 at the same time, and did we get an answer to the FiO2 question?

DR. CANTRILL: The majority of the patients that are ambulatory, their FiO2 was 0.2.

So it's a much smaller number that are on -- are on supplemental oxygen when they present to the ED, and again that's when -- that's when your vital signs are taken.

The nurses are usually pretty good about documenting your -- any supplemental oxygen the patient is on at the point of presentation.

CO-CHAIR WEISS: So that, let me understand, clinically, I guess the question is specification wise, what does it say about

FiO2 and the need for FiO2 --

DR. ANTMAN: So in specifications for the measure, the FiO2 is not required to meet this measure. If I may add, Dr. Weiss, as far as the evidence for the measure, in the American Thoracic Society guideline that we used as a reference for this measure, there is the following statement:

"For those patients with chronic heart or lung disease, the assessment of oxygenation by pulse oximetry will help identify the need for hospitalization."

So clearly it's recommended in the ambulatory setting.

MEMBER EDELMAN: Well you have to read the whole sentence. It's people with chronic cardiopulmonary disease.

So what is the evidence that in a primary care setting, a family physician, that failure to do oximetry associated with diagnosis of pneumonia, leads to a poorer outcome?

DR. ANTMAN: Right, so looking among my colleagues here, it doesn't appear that we have any of that --

CO-CHAIR WEISS: Just another of the 15-minute marks on this measure.

MEMBER YEALY: I guess I would just say the most commonplace presentation for acute community-acquired pneumonia in the emergency department, the evidence is absolutely clear there that you can't restratify absent oxygenation, and while we think there are many other folks who are given a more colloquial diagnosis of community-acquired pneumonia in a different setting, there is nothing to refute this and there is no structure --

MEMBER EDELMAN: I was having no problem with this when it was confined to the emergency department.

MEMBER YEALY: No, I understand.

MEMBER EDELMAN: Now it's going to be used to whack a lot of GPs in the head and

I think we need some evidence.

MEMBER YEALY: Or change their =-= or change their diagnosis if they are not going to truly seek --

MEMBER EDELMAN: So this is --

MEMBER YEALY: Let me finish please for as second. To actually diagnose acute lower respiratory tract infection since they are not going to get the radiograph either.

So they couldn't -- they couldn't have really diagnosed -- they may have suspected and I'm not -- that's a whole different conversation.

MEMBER EDELMAN: So now we are addressing -- we are not addressing care. We are addressing the upcoding.

CO-CHAIR WEISS: Well there's actually, it may be -- let me just suggest that we are entering into the territory of performance characteristics of the existing instruments and the pre-test probability

associated with what patients are coming in, in terms of underlying prevalence.

So, the emergency -- all the performance characteristics of using these and building these instruments, these tools, have been based upon presumptive emergency medicine comers, as opposed to what would come into a primary care.

So there's probably different performance characteristics that are associated with the need for this, but we don't know that.

All that said and done, we are left with the measure here, which has been specified that it would include both emergency department and ambulatory diagnosis, and one way to look at this is exactly what we've seen, which is that this would elevate care because it would help the outpatient care that is happening outside the emergency room for diagnosis treatment look a lot more like emergency room care, and then there's the

other one saying -- other balance to this saying well, but this may be actually a different treatment algorithm would be viewed here, a different process of clinical diagnosis and treatment.

I don't know where the answer is here but I think those are the issues that you are raising. Do I have them right in terms of what's being propagated? Okay.

MEMBER STEMPLE: You said would we be elevating the level of care in a PCP's office and I don't think we heard there's evidence, just because you do a pulse ox, I didn't hear that we had evidence that that elevates care outside of the chronic population, which is not what this measure is saying. This is saying all coming. So I think you misspoke in a way, because you said we would be elevating the level of care. I don't know that there's evidence to say that we are elevating the level of care.

CO-CHAIR WEISS: That's great.

Let me clarify my statement, which was we would elevate the level of care to provide the -- by providing O2 sat, you could actually do a better risk stratification of the persons who come into your office with pneumonia, similarly to what they would get in an ED.

And that's the level of assessment elevation that you'd get, is the presumption here. I'm not arguing for it or against it. I'm just speaking to what I think I'm hearing are the issues. Trude.

MEMBER HAECKER: Are there diagnostic issues as well? I'm not going to do an x-ray in your office because I'm 20 minutes away from the radiology suite when I'm down the hall from a portable chest x-ray in the emergency room. And that does have to be factored in as well?

CO-CHAIR WEISS: Yes. Okay. Have we got all the issues on the table? I mean, everyone is going to have to vote for themselves here, but --

MEMBER STEMPLE: I'm sorry, any idea of what percent of this population is ER versus ambulatory care treated by -- or non-ER outpatient treated?

CO-CHAIR WEISS: Do we have any sense from our measure developers of what proportion of CAP is diagnosed like the NAMCS or something, you know, the ambulatory care surveys that show what proportion of --

DR. ANTMAN: I don't think we have that separated out.

CO-CHAIR WEISS: Okay, so we don't have that information. Okay. Do we have all the issues on the table? Can we go to a vote? Would you be? I mean for better or for worse, we can figure this one out?

Has it met the criteria, yes, no, insufficient evidence?

(Pause for voting)

CO-CHAIR WEISS: I'll be curious to see what this shows. Oh come on, the suspense is killing is. Okay. Everybody

press again. Doesn't want to cough up this fur ball. Okay let's try one more -- a third time.

Where are we at here? Okay. We are going to time ourselves out because we are not getting the 20th here. Someone is not wanting to vote or someone's battery is running out. Three, two, okay. Based upon 19 votes here's what we've got.

Ooh. A bit of a surprise. Okay. Five yes, two no and 12 insufficient evidence. So after a long discussion, a labored discussion, we have to say no to this measure. The -- I don't think we need to review the specifics. You've got them down on record. We've got the measure developers here.

I think there's a lot of intrigue about this measure. But the unanswered question seems to be to -- seems to be the higher order of the day. Okay?

MEMBER RHEW: Well I also -- I think the limiting it to the ED would probably

-- you might have a different result.

CO-CHAIR WEISS: So you're not saying a big affirmative to the committee on that. Did you want to say one other closing comment as the measure developer? Steve, do you want to?

CO-CHAIR GROSSBART: Well, I was just going to say limiting it to the ED may remove the ambiguity but in the inpatient setting, and all those patients are coming through the ED, we are at 100 percent in the bottom decile, so it -- there may not be a gap in the ED.

DR. ANTMAN: So if I may, with regard to the setting, and again, apologies that we apparently created some confusion as to the setting.

I look to my colleagues to correct this statement if I have this wrong, but I believe that the pneumonia measures, the PCPI's suite of pneumonia measures, were initially created for the variety of non-

inpatient settings that we have been describing here -- ambulatory, physician office, other ambulatory settings including the ED.

When we convened a group of emergency physicians like Dr. Cantrill to look at measures that would be particularly useful in the ED setting, we adapted many of those measures for the ED setting.

So we have a separate set of pneumonia measures that are specified separately for the ED setting. Clearly there is -- there are issues with the evidence for the ambulatory settings other than the ED, so that's valuable for us to hear back.

If I may add just one other quick note and that is that the confusion related to the statement about there being measures in Hospital Compare, that is not a note that we, the PCPI staff, inserted into any of our submissions. I think that was a clarification that -- okay. Okay. Thank you.

CO-CHAIR WEISS: Because they were confused and so that's where they went for the data.

MEMBER PELLICONE: Can I ask one more, one real quick clarification on the measure just even though we didn't get to it? Is it a requirement that there be documentation, that the clinician saw the O2 saturation, and not just a printout of the fifth vital sign?

CO-CHAIR WEISS: Something to consider. No need to respond to that. Unless you've got a quick response. Otherwise let's continue. I think the answer is --

MEMBER JEWELL: So for me, listening to this whole conversation, the problem I had was the same problem I had with the dyspnea measure. It sounds like we are talking about diagnostic tests and measures that may well have very important clinical meaningfulness for specific populations in the outpatient or ambulatory setting, or the home

health setting, and if that's true, then those are the groups we should be targeting, not this gunshot approach where we might capture everyone and that gives people more opportunity to participate.

CO-CHAIR WEISS: Okay good. Let's go to the third of these PCPI measures, which would be the -- oh sorry. The vital signs for community -- for CAP.

Uh-oh. Okay. Dr. Yealy.

MEMBER YEALY: Well, I think we are going to have the same basic conversation again. It is, it runs parallel to oxygenation in that the vital signs are central to any restratification that you will do in a community-acquired pneumonia, no matter how you diagnose it, and will impact upon your decision to treat as an in- or outpatient and what type of coverage you'll do, because as you move up the severity scale, we often broaden coverage and we know that in sicker people broadened coverage is associated with

better outcomes.

So there's a variety of dominos that have to fall, but this is a pretty basic one. The rub here is, is we don't actually ask you to integrate any of those particular -- we don't tell you how to integrate them or ask you to show us how you did. Did you use PSI or CURB or SCAP or pick whatever tool you like. We just say please make sure they're measured somewhere.

So there is a behavioral gap in there that's not measured. Nonetheless, although I thought oxygenation was going to be on face value a no-brainer, this is really at all levels the exact same conversation.

MEMBER EDELMAN: May I ask a question? The denominator is bacterial pneumonia. Am I correct?

MEMBER YEALY: I believe so.

MEMBER EDELMAN: In community-acquired pneumonia, what percentage of pneumonias are documented by laboratory

methods to be the bacteria?

MEMBER YEALY: You could just say in pneumonia in general, what's the frequency of diagnosis?

MEMBER EDELMAN: Oh no, but it says bacterial.

MEMBER YEALY: I could just -- you could ask in general the number of times a pathogen is identified, whether it's community or non-community is actually probably in the order of 30 percent.

MEMBER EDELMAN: So why --

MEMBER YEALY: That doesn't mean the others are not, it's just an --

MEMBER EDELMAN: So why limit the denominator to bacterial? I mean -- no, but I'm asking whoever the people who wrote this, why on earth are you focusing on bacterial?

CO-CHAIR WEISS: So I'm going to be mindful of time here, because we're -- I'm quickly becoming a very bad timekeeper, even with our 15-minute marker here.

Let's -- with that bit of general background, let's go right into anything else you'd like to say, Don, about impact, performance gap --

MEMBER YEALY: No, I mean, we could go through this all again, but it is the exact same conversation one more time.

CO-CHAIR WEISS: Is that what I'm hearing from -- is that -- okay, well I'm hearing at least enough people saying that they are not certain of that that we need to go through it.

MEMBER YEALY: I guess the only difference would be most of us wouldn't wonder about whether a community physician or an office had measured the rest of the vital signs. That's the only thing that --

CO-CHAIR WEISS: So it sounds like we can go through, we need to go through this, but at least impact and performance gap, let's -- and I'm shepherding us through a formal process because we need to go through it.

Anything else you'd like to say about impact or vital signs?

MEMBER YEALY: Yes.

CO-CHAIR WEISS: Anything you'd like to say about performance gap and what at least is --

MEMBER YEALY: So I turn to the folks who have been -- who have proposed this and followed it, and let us know what the performance gaps are on vital signs, the most recent data. We did not have that at the call.

Again, I know from even RCT work, as frightening as this sounds, how often they are not completed.

DR. CANTRILL: The 2008 PQRS data gives a performance gap of 22.3 percent.

CO-CHAIR WEISS: I mean that seems like -- one out of five patients don't get vital signs?

MEMBER YEALY: Don't get complete vital signs.

CO-CHAIR WEISS: Complete vital signs. And this is because we are now mixing -- not because, but we know that we are mixing emergency room and ambulatory --

MEMBER YEALY: Mixing all ambulatory types and essentially three different sets -- three different variables.

CO-CHAIR WEISS: Okay. So in terms of performance gap, we see about 22 percent is the number, in this 2009 data we think?

And then we have the final piece of evidence supporting this.

MEMBER YEALY: Again, it comes back to these are the cornerstones, no matter which rule you use, the only question is whether you use two or three of these in your assessment of severity, which will drive almost everything downstream, including whether you will just decide to send someone home on oral therapy or refer them on elsewhere or send them to the emergency

department.

There's really no way you can make that decision without having incorporated this evidence.

MEMBER STEMPLE: And Kevin, we have heard O2 sat is the fifth vital sign, is O2 sat one of the vital signs here?

CO-CHAIR WEISS: That was more conceptual. They are trying to make it that.

MEMBER STEMPLE: Okay, I just wanted to make sure. Temperature, respiratory and blood pressure.

CO-CHAIR WEISS: The pulse, respiratory, blood pressure. Okay. Dianne and then Norman.

MEMBER JEWELL: So forgive me if everybody else understands this but me, but does it matter that we are talking about this relative to bacterial pneumonia as opposed to all the pneumonias? Okay.

MEMBER LEVY: Well, it shouldn't. There's nothing about bacterial that's --

MEMBER JEWELL: That makes this extra special. So --

MEMBER LEVY: That's what you were asking.

MEMBER JEWELL: So the notion. Yes. Okay. I just wanted to -- because I didn't actually hear the answer to your question, so that's why I was just trying to clarify. Thank you.

MEMBER YEALY: If I was writing it, I would have just excluded that word. I would have just said acute community-acquired pneumonia and leave it at that.

CO-CHAIR WEISS: Can I ask from a developer's standpoint, is that a clarification opportunity, or is that a new measure, I mean is that changing the measure specifications?

DR. ANTMAN: And I'm sorry, do you mean the bacterial specification?

CO-CHAIR WEISS: Yes.

DR. ANTMAN: So although the

guideline citations don't say this specifically, I believe, and we're checking, I believe that the guidelines that we reference were specific to bacterial pneumonia, but that's something we can certainly verify.

If I can address another question that came up. In our specifications, we do state that the vital signs include temperature, pulse, respiratory rate and blood pressure.

It does not include O2 sat.

CO-CHAIR WEISS: Mitchell.

MEMBER LEVY: Yes, the -- I worry -- the numerator, in the gap analysis we are quoting, the numerator here says documented and reviewed, and I'm not sure I understand that.

Does that mean it has to be documented in a chart somewhere that someone -- not only that --

(Alarm sounds)

MEMBER LEVY: the vital signs have

been done, but they have been reviewed by the clinician? Because that's -- is that what is being tracked when you -- we have quoted the gap analysis data? Those are two very different things, and I definitely want clarification of that.

DR. ANTMAN: Well, the intent is that it will have been reviewed by the clinician, yes. And that -- and going back to an earlier question, that was the intent for the O2 sat measure as well. I think someone asked, does that mean that the physician actually looked at the results, and the answer is yes. We added the words "and reviewed" specifically, because that's the intent.

MEMBER LEVY: But that means, so that's either self-attestation, or I mean, that -- how you measure that, that's, that makes that measure very difficult to -- that metric very difficult to measure.

I mean, almost impossible unless somebody is reviewing charts. That's why I

ask, because as written there, then that's a metric that really is not -- is almost unusable but we are quoting gap analysis, so I am wondering if now we are conflating these two -- these two approaches.

DR. RALLINS: I'd like to make one comment. I'm Marjorie Rallins and I work with the specifications team, and we are also working on developing this specification for an electronic data source that it is anticipated that you can capture additional nuances such as documented and reviewed differently than you can if you are reviewing claims.

So I'd like for us to be mindful of that when we are having our discussion.

CO-CHAIR WEISS: That's in the developmental phase, or have you implemented it and tested some of this or --

DR. RALLINS: We are developing our specifications, I don't believe those have been tested or --

CO-CHAIR WEISS: So that's something for us to look towards in the future more, so then --

DR. RALLINS: Sure. Sure.

CO-CHAIR WEISS: Okay. Good. Excellent. It sounds great. Norman.

MEMBER EDELMAN: Very briefly. This is not like the oxygen. This is the opposite of the oxygen. So it's perfectly reasonable to ask a GP to take vital signs. But the impact is trivial if it's limited to documented bacterial pneumonia, because that almost never happens in the ambulatory setting.

CO-CHAIR WEISS: Comment to our measure developers. So your thoughts on what we just heard from Norman about the fact that this bacterial pneumonia specification is -- sounds like it's embedded because that's where the literature was and it went from the emergency room, which is where -- so going backward, by keeping the bacteria, which is

something you don't see quite commonly in an ambulatory setting because of the lack of diagnosis and the lack of sending cultures and all that kind of stuff. Interesting. Okay.

MEMBER YEALY: The data aren't specific to bacterial and I suspect you are not just pulling bacterial -- because in fact it would be almost impossible to construct such a cohort.

So this again I see as an opportunity for clarification, not change of the guideline. I just think that for whatever reason, that word ended up in there infortuitously.

CO-CHAIR WEISS: Okay.

DR. ANTMAN: So I apologize that we don't have a definitive answer to that but we will certainly find one.

CO-CHAIR WEISS: Peter, and then we are going to go to vote.

MEMBER ALMENOFF: I have just one last comment for the developers. It wouldn't

be a bad idea if you are going to do a lot of ambulatory care measures to maybe have the American College of Family Practice or another group out of an ED. You know, ED is going to be good at ED, but you know, we might have gotten another perspective from the family practitioners regarding oximetry in an office, so it wouldn't be a bad idea to try to get some perspectives of the groups that really -- because at least here it shows me ED and some other group, but it doesn't show anything regarding family practice or any of the people who really do most of the ambulatory care work.

DR. ANTMAN: And forgive me doctor, I'm sorry, are you looking at the makeup of our development group?

MEMBER ALMENOFF: Right.

DR. ANTMAN: Okay.

MEMBER ALMENOFF: Because it would have eliminated maybe the issue about oxygenation, because they might have said we

do that all the time. I mean, I don't know.

CO-CHAIR WEISS: Let me, if I can, because we have to get to vote, just log that as a thought, take it offline if you want to, to find out more detail.

But let's vote. Impact. One high, two moderate, three low, four insufficient evidence.

(Pause for voting)

CO-CHAIR WEISS: Seventeen, 19. Okay who is that, there we go, 10 high, 7 moderate, 2 low and 1 insufficient evidence. Next let's go to performance gap. High, moderate, low, insufficient evidence.

(Pause for voting)

CO-CHAIR WEISS: Okay, 7 high, 11 moderate, no low and 2 insufficient evidence. Okay, let's continue on to -- is there enough evidence in your mind to support this going forward? Yes, no or insufficient.

(Pause for voting)

CO-CHAIR WEISS: We are at, what

18, 19? Okay. There we go. And the answer is yes, 16, 1 no, and 3 insufficient evidence.

Okay, that means we talk about reliability and validity. Don?

MEMBER YEALY: And I think we have already touched upon the reliability issue, about the chasm between documentation of the vital sign and, or the measuring of the vital sign and someone's actually knowledge of it, where that can offer the appearance of lack of integration of this information, and that was the only concern that I recall before, and it didn't outweigh the positive recommendation.

CO-CHAIR WEISS: Let's go on to -- well first ask anyone on the working group have any thoughts about reliability? And how about the committee as a whole?

(No response)

CO-CHAIR WEISS: Let's shift to validity. Let's shift to the validity discussion. Yes. Not there yet.

MEMBER YEALY: Let me pull up my

notes, hang on one second. Again, there wasn't a lot of concern about the validity issue, again, we didn't have much conversation.

Again it cycles back to the issue of knowing them and integrating them fully are two separate things and we can't -- and we don't have a way, unless we -- unless a brand new criteria was developed that said use and then gave you a menu of that and that's a completely different, that's not on the table right now.

CO-CHAIR WEISS: Okay. From the working group? From the table at large?

(No response)

CO-CHAIR WEISS: We've talked this one out. Okay. So let's now vote for reliability. Reliability, one, two, three, high, moderate, low or insufficient evidence.

(Pause for voting)

CO-CHAIR WEISS: Here we go, 10 say high, 8 say moderate, 2 say low and no for

insufficient. Let's go to validity testing. High, moderate, low and insufficient. Please vote.

(Pause for voting)

CO-CHAIR WEISS: It is in there. Correct. Are you seeding thoughts in our head?

Here we go. So seven high, nine moderate, three low and one insufficient. And so it goes on to usability and feasibility. So --

MEMBER YEALY: Again, we didn't see any concerns about this particular part, as it -- as it connected to outside understanding.

CO-CHAIR WEISS: Workgroup? Total group?

(No response)

CO-CHAIR WEISS: Vote one, two, three, high, moderate, low, insufficient information.

(Pause for voting)

CO-CHAIR WEISS: Thirteen high, five moderate, two low and none for insufficient information. Let's go to the final item, element, feasibility. How feasible is it to collect this data? I guess that means --

MEMBER YEALY: Again, this is -- we have talked about it before. Collecting the documentation of the vital signs is not a challenge. Determining whether or not it's been integrated is a whole separate issue but it's not part of the measurement.

CO-CHAIR WEISS: Okay. It is. Review is part of it.

MEMBER YEALY: No, I mean, how it was -- how you integrated that information into your decision-making is what I'm saying.

CO-CHAIR WEISS: Oh, okay.

MEMBER YEALY: There's no -- you know the implicit hook of collecting the vital signs is that you will use them appropriately to make a decision and that is not what is

being asked here.

CO-CHAIR WEISS: This is a documentation measure. Full stop. Okay. So --

MEMBER YEALY: There's a leap of faith of that having them will make you act appropriately.

CO-CHAIR WEISS: And then any comment from the group, otherwise we'll vote. I guess we're voting. Good.

(Pause for voting)

CO-CHAIR WEISS: There you go. We got all 20 there. Good. And high, moderate, low, oh sorry -- high nine, moderate seven, low three and insufficient one. It's getting kind of like -- it's getting hypnotic.

Okay, here it is. This is the overall shall we endorse, send this forward for endorsement?

Yes, no. One, two.

(Pause for voting)

CO-CHAIR WEISS: Seventeen say

yes, three say no. It does raise a question in my mind's eye at least. I remember when ophthalmology had the measure that you needed to look at -- do an eye exam before doing a cataract extraction, I mean the vital signs before you work on pneumonia feels like we are hitting a real lowball in measurement here.

But it seems to pass through the process well. Good. Let's -- what was that? Oh yes. So, but it seems like there's more technically getting through there. But okay, 1895.

And we -- once again we have to be mindful we have got our colleagues from CMS on the phone and we want to thank our colleagues from CMS and hope they will be us for a little bit longer.

Assessment of the mental status for community-acquired bacterial pneumonia. We don't have to talk about the word bacterial again, because we have heard that, and Christy, you're up at bat here.

MEMBER WHETSELL: Well, I think everything comes across that we have discussed in the last two, about bacterial, about what environment is this being collected in, and things like that.

When we were on the conference call as a team, we kind of felt this was a no-brainer, you should do a mental exam on a patient, and we kind of sailed through it.

CO-CHAIR WEISS: Okay. So in terms of its impact -- what was that? This is same as before. It's --

MEMBER WHETSELL: Same as before.

MEMBER RHEW: Quick question. Steve, earlier you said that this is actually not mental status exam but this is confusion. Is -- do we need to change that then?

CO-CHAIR WEISS: Well that would be changing the whole specification.

DR. CANTRILL: I don't think so but that is the data point that we are after because that is used in some of the algorithms

to determine when should the patient be admitted and should they be admitted to an ICU?

So it's -- mental status is a routine part of a physical exam.

MEMBER ALMENOFF: We're talking about all settings though, so --

DR. CANTRILL: I'm sorry?

MEMBER ALMENOFF: We're talking about all settings. So usually --

DR. CANTRILL: I would maintain that mental status is a routine part of a physical exam.

MEMBER EDELMAN: So again in the family doctor's office, you would require his note to indicate that he had done a mental status exam? Is that correct? Thank you.

CO-CHAIR WEISS: At least oriented times three.

DR. CANTRILL: Are outpatient offices included in this, even if they are not part of a hospital?

MEMBER ALMENOFF: Correct.

DR. CANTRILL: So that would be true.

CO-CHAIR WEISS: Okay, so impact, any other comments from the group on impact? If not, let's think about performance gap. What are we seeing as performance gap currently in terms of --

MEMBER WHETSELL: Looking at their data that they talked about, I think they said there was a 20 percent gap, I'm sorry, 19.42 percent gap.

CO-CHAIR WEISS: Very similar to the vital signs. Suspiciously similar to the vital signs. But good. Okay. And then the third element is the evidence.

MEMBER WHETSELL: I think the concern there was that there is variation in how mental status can be evaluated.

CO-CHAIR WEISS: Okay.

MEMBER WHETSELL: And/or reported.

CO-CHAIR WEISS: Did that

variation concern anybody in terms of how it was being specified?

MEMBER WHETSELL: To my team, I don't recall us having that discussion.

CO-CHAIR WEISS: Okay. Very good. Then let's go --

MEMBER EDELMAN: I'm sorry. I have a question. And there is evidence, presumably, that failure to assess mental status leads to inappropriate clinical outcomes?

CO-CHAIR WEISS: Does or can?

MEMBER EDELMAN: Well, I mean, there should, since you are taking vital signs, and doing oximetry and doing all kinds of other things, it might be something that doesn't add to the clinical decision-making.

MEMBER YEALY: So the single biggest point score in the pneumonia severity index is altered sensorium, single biggest change, actually outside of age.

That would outweigh almost -- it

goes neck and neck with hypotension so --

MEMBER EDELMAN: So that would lead to sending a patient to the floor rather than the ICU, is that correct?

MEMBER YEALY: In theory, if you hadn't assessed it --

MEMBER EDELMAN: In theory.

MEMBER YEALY: Or sending home instead of admitting to the hospital or making -- you wouldn't be able to fully, and the same happens for CURB and SCAP. They all use some assessment.

And what literature there is, does not suggest that differing tools leave you in dramatically different spots, whether you use sensorium, confusion, Glasgow Coma Scale which was never intended for this, that it -- that some look ends up getting you where you need to be.

MEMBER ALMENOFF: But isn't most of that literature in the ER setting?

MEMBER YEALY: It's actually ER in

inpatient setting. Again --

MEMBER ALMENOFF: And so --

MEMBER YEALY: I just don't, there's not --

MEMBER ALMENOFF: That's why we keep extrapolating data from one setting and putting it in another setting, and I'm just not --

MEMBER YEALY: I can't speak and I don't know of large cohorts in the ambulatory setting. I just don't know of them.

MEMBER ALMENOFF: Right.

CO-CHAIR WEISS: So, with that in mind are we ready to -- any other questions from the group as a whole on these three constructs? If not, let's vote on them.

Impact.

(Pause for voting)

CO-CHAIR WEISS: Okay, let's try again. Please. Press your buttons again. There we go. Perfect. So high eight, moderate eight, low one, insufficient evidence

one. Next would be the gap which we heard was 20 some -- 19 point something percent.

MEMBER WHETSELL: 19.42.

CO-CHAIR WEISS: Okay.

(Pause for voting)

CO-CHAIR WEISS: Okay we are voting on performance gap. Good. All the ones are in. So 6 high, 13 moderate, no low, no insufficient. Next is the evidence.

(Pause for voting)

CO-CHAIR WEISS: Is the evidence clear? Yes, no or insufficient. Okay. Wow, so 14 yes and 5 insufficient. So it moves forward.

So let's talk about reliability and validity. Christy it comes back to you.

MEMBER WHETSELL: I think in reliability the discussion we had had was again, variation of a tool used can impact what we see. Validity, we didn't have a discussion.

CO-CHAIR WEISS: So let's ask. Is

there concerns about validity that we would need to look at here? This goes to the group as a whole. Rubin.

MEMBER COHEN: I'm just wondering. So just documenting that the patient is confused is adequate?

CO-CHAIR WEISS: That's what we are hearing.

MEMBER COHEN: Or do you have like loss of recent memory, or orientation? How extensive does this have to be?

The patient has a fever, he's confused. That's adequate to assess -- just.

DR. CANTRILL: I think this would be passed by listing any component of a mental status.

MEMBER COHEN: Any component?

DR. CANTRILL: Yes.

MEMBER COHEN: Okay.

CO-CHAIR WEISS: Okay. Any other questions, thoughts? Let's vote on reliability and validity. One, two, three,

high, moderate, low on reliability, four if you feel it's insufficient information.

(Pause for voting)

CO-CHAIR WEISS: Get those fingers moving, a little afternoon exercise on the fingers. Almost there. If you vote we can save 40 seconds here. Press again.

Look everybody, point to Jessica, let's do it again one more time. Save 30 seconds. Come on, you can do it. There we go. Okay.

So 5 high, 11 moderate, 2 low and 1 insufficient evidence. Let's go to validity. Yes.

MEMBER JEWELL: So there wouldn't be exclusions for people who already have documented problems like dementia or other cognitive decline?

CO-CHAIR WEISS: How is that handled? For a person who is -- well you are still assessing it. The question is, does it contribute much. One might argue that maybe

people who have altered mental status even though their pneumonia is not bad, may have a hard time with compliance, particularly in outpatient. So, but we are just talking about documentation here, did it document, and whether that relates to outcome.

MEMBER JEWELL: Okay, so and I understand we are not talking about how you use the information, I guess, well, except we have been talking about risk stratification as the evidence, so I guess that just --

CO-CHAIR WEISS: Correct. That's -- you're right.

MEMBER JEWELL: That for me is a disconnect, but okay.

CO-CHAIR WEISS: Well one would assume that a person who has dementia may be at a higher risk and again, it may be for reasons that are not related to the bacterial pneumonia per se, but maybe to compliance or ability to express the need for anything to additional therapy.

MEMBER YEALY: Again, I didn't write the criteria that existed before I came here, but as a PSI author, any change in mental status, whether it was new or old, is a bad thing.

MEMBER JEWELL: And so that's the -- because it was change of mental status that I heard, and so I was thinking it meant acute change, not longstanding change also. Thank you.

CO-CHAIR WEISS: So it's just altered mental status. Good, so let's vote on validity. One, two, three, high, moderate, low and four for insufficient evidence.

You guys are going to be happy to get rid of me with two more measures.

(Pause for voting)

CO-CHAIR WEISS: Okay. Six high, 12 moderate validity, one low and no insufficient, so it passes through. Let's go to usability.

Back to you Christy. Usability

and feasibility.

MEMBER WHETSELL: Again we thought that this was just a no-brainer, that it would be highly useful and feasibly easy to obtain.

CO-CHAIR WEISS: I assume that the review issue still was standing about how you know that it was -- that it was actually reviewed. Is that -- review is part of this as well? It's document and review, or is it just document?

DR. CANTRILL: It's nominally documented by the physician. So since he is the decision maker here, that is implied that if he evaluated the patient for that, that in fact that would be part of his decision-making process.

CO-CHAIR WEISS: Okay. Very good. Questions, thoughts, comments around the room?

(No response)

CO-CHAIR WEISS: Okay, then let's vote first on usability. One, two, three, and

high, moderate, low, four is insufficient.

(Pause for voting)

CO-CHAIR WEISS: Let's try again, see if we can move this timeframe a little faster. There we go. All set.

Seven say high, 12 moderate, no low, no insufficient, okay. Let's go from usability to feasibility. High, moderate, low and insufficient.

(Pause for voting)

CO-CHAIR WEISS: Okay if everyone can try again so we can try and speed up the clock a little bit, that would be wonderful. There we go. Good.

Six say high, much more moderate -- - 12, one low, and no insufficient. Let's go to the final, summative vote. Yes this should be moved on for endorsement, and no way. One or two.

(Pause for voting)

CO-CHAIR WEISS: I think we got 18. Let's try again. Push everybody. It's a

high incentive to speed us along here. One more time everybody, come on, we can do it. There we go. Okay. Nineteen yes, unanimous. Move it forward.

Again, this is a documentation measure. It does feel a little bit lowball. But it sounds like the community is not doing it so good.

We are down to two measures to complete this section of process measures for pneumonia. Both are CMS measures. I'm just wondering if it's -- we are running a little bit late. It feels like a break would be required, just biologically.

Finish, want to push through these, or do you want to -- do people need to take a quick break because of human dimensions here, a biologic moment?

Let's -- who is on the phone with us? Dale?

DR. BRATZLER: Dale Bratzler.

CO-CHAIR WEISS: Dale, would you

mind if we took like a 5, 7, 10 minute break just so we can get -- we haven't had a break since lunch? Would that be painful to you? Or not?

DR. BRATZLER: I'm pushing up against another meeting, but I can certainly wait.

CO-CHAIR WEISS: Okay, well let's just do it. Okay. Let's just do it. So if you are going to -- please, if anyone needs to step out the room, please do so, but do so really as expeditiously as you can.

Let's jump right in. Measure 0147.

DR. BRATZLER: I think I can give a very quick overview, mainly because you have already talked about 0147 largely, it's not that much different from this AMA measure.

It's the initial antibiotic selection for community-acquired pneumonia is a measure that focuses only on those patients admitted to inpatient status. We do have a --

so as you heard earlier, the denominator is defined by a patient that has a discharge diagnosis of pneumonia.

However, we have a data element that says did the emergency department or the initial admitting physician make a diagnosis, a clinical diagnosis of pneumonia.

So a pneumonia that is diagnosed subsequent, during the hospital stay, is not included in the measure. And it's -- the performance measure is based on the IDSA in the American Thoracic Society guidelines from 2007.

However the measure is continuously updated. We meet every three months and if you look at the performance measures classifications, the measure has been substantially updated since 2007 because there are new antibiotics on the market, and we do meet with the guideline panel every three months to talk about updates.

The second performance measure,

0148, focuses on only those patients who have an -- who have a blood culture drawn while they are in the emergency department.

In other words, if the patient is not in the emergency department, if the clinician decides not to order a blood culture --

(Alarm sounds)

DR. BRATZLER: they are not in the denominator for the measure. The measure simply looks at if the emergency department physician decides to order a blood culture, do they draw -- have the blood culture drawn before the antibiotics are given.

Once the decision to admit the patient is made, and it's documented in the chart, then the patient is no longer eligible for this measure. It only focuses on those patients in emergency departments, when and only for those patients, completely at the discretion of the ED physician to decide whether a blood culture is needed or not.

CO-CHAIR WEISS: Excellent. Thank you so much. Before we begin the specific discussion, any from the table, any general questions to measure developer?

(No response)

CO-CHAIR WEISS: If not, then Dave.

MEMBER RHEW: Sure, thanks Dale, and in fact this is very similar to our prior discussion on the 0096 so what we will really just focus on are the key differences. The key difference certainly rationale wise, I mean this is clearly an important initiative, we also know that from the most recent Hospital Compare, the current adherence rate is 94 percent.

And as per prior discussions, the evidence is very strong, much more so in the inpatient side and the severely ill patients, as opposed to the lesser sick.

So I think those are kind of the three things in terms of the rationale, the

gap and the evidence, but it's really pretty much the same discussion that we had previously.

CO-CHAIR WEISS: Okay, and let's ask the workgroup, would you like to add any thoughts to what Dave has suggested?

(No response)

CO-CHAIR WEISS: Okay. And from around the table. Questions or thoughts with relationship to impact, gap or evidence?

(No response)

CO-CHAIR WEISS: Okay. Everyone still with me? We have an -- I just want to make sure that we are not into a -- we are all here, yes? Yes? This is it for this time of day, huh?

Okay, let's vote. Impact.

(Pause for voting)

CO-CHAIR WEISS: Thirteen high, six moderate, no low and no insufficient evidence. Let's move on to the gap. Did we hear the performance gap? I don't remember.

Ninety four percent?

MEMBER RHEW: Ninety four percent.

CO-CHAIR WEISS: Do we want to say anything about that or --

MEMBER RHEW: Again, that there is a gap. We -- it could probably be improved upon, and again, especially in the inpatient -- I guess the one thought that we did have that we know is not currently in there, but if there was an ability to tease out ICU versus non-ICU, that was something that we thought could be very helpful because the impact is much stronger in ICU.

CO-CHAIR WEISS: Is there much variability in that 94 percent, or is it really just --

MEMBER RHEW: That's what we'd like to know, and that's -- that's I think where we could really better understand if the opportunity is much greater than what we really believe it to be.

CO-CHAIR WEISS: Do we have a --

information from Dale, do you have anything with relationship to variability with -- because that's a very high success rate, 94.

DR. BRATZLER: Yes, so we did forward information earlier about the performance measure and the, you know, the greatest opportunity for improvement is still you know, in the intensive care unit setting, where rates of performance are much lower for those patients, the sickest patients and get them into the ICU.

But I think we did provide a nice distribution of the number of hospitals that pass or fail the measure. There is still substantial opportunity for improvement but particularly for the ICU population.

CO-CHAIR WEISS: Okay. Great. Any other thoughts or comments related to gap? Otherwise let's vote.

(Pause for voting)

CO-CHAIR WEISS: Okay. Here we go, eight high, eight moderate, two low, and

one insufficient evidence. Let's go to the final of these three, evidence. Yes there is adequate evidence, two, no there is not, and three is insufficient evidence.

(Pause for voting)

CO-CHAIR WEISS: Let's press again if we could, just to see if we can boost this along. There we go. All set, 17 say yes, 2 say no, no insufficient. Let's move on to reliability and validity.

Dave?

MEMBER RHEW: Sure, with regards to reliability, I mean, I think Dale you have done a really nice job in terms of including all the key specifications, but the one thing that perhaps would be nice to include, you mentioned this, but some reference to what the antibiotics were or how, when you update them, how we would know, and you made reference to the IDSA/ATS guidelines.

But some reference to that would be helpful in the document, recognizing that

we are not relying on the 2007 guidelines per se, but this is an ongoing area where we could perhaps get, you know, maybe tap into so we know what the specifications are.

CO-CHAIR WEISS: Okay. Comments from the rest of the workgroup. Dale?

DR. BRATZLER: No, I don't have anything to add. We -- that's why I mentioned, we do update the performance metric. If you -- the manual gets updated twice a year but the panel meets every three months.

CO-CHAIR WEISS: Okay, the rest of the workgroup, or the rest of the table?

(No response)

CO-CHAIR WEISS: Okay. Validity? Where are we at in validity?

MEMBER RHEW: Again, this is one where we -- well, this has been implemented, it's been tested, it's been pulled out from the EHR, I mean we thought this is a highly valid, highly reliable metric.

The caveat of course being you know, what we mentioned before, this is post -- you know, this is all retrospective as opposed to prospectively collected and that's a challenge that we all face.

CO-CHAIR WEISS: Okay. Very good. Any comments from the table?

(No response)

CO-CHAIR WEISS: Then let's vote. High, moderate, low or insufficient, on reliability.

(Pause for voting)

CO-CHAIR WEISS: Again we are going to vote again. Done. Okay good, 17 high, 2 moderate, no low and no insufficient. Let's go on to validity.

MEMBER ALMENOFF: And I guess this is for maybe it's CMS. You know you have a list of all the -- in table 3.1, all the, I guess, diagnoses of pneumonia.

Is this the entire list you are going to use, because you have a lot of things

on here that I wouldn't consider CAP. So I'm just kind of curious, because you know, with the -- the CMS Hospital Compare measure, people kept -- CMS kept taking out different diagnostic codes because they weren't correct, and it just seems like you have almost everything and the kitchen sink in the diagnoses here.

So would anybody be able to address that?

CO-CHAIR WEISS: Dale, did you hear that comment?

DR. BRATZLER: Sorry, I did not.

MEMBER ALMENOFF: So let me repeat that again. Under the table 3.1, with the pneumonias -- can you hear me?

DR. BRATZLER: Yes, table 3.1. I'm just not sure what you are referring to, but --

MEMBER ALMENOFF: You have got a list of almost every organism on earth on the list and a lot of them are not associated with

community-acquired pneumonia. So just sort of wondering why they are all on this list. Shouldn't it be selective organisms that we would be thinking about regarding CAP and not gram-negative organisms, and MRSA septicemia and Pseudomonas. I always thought those were not CAP-related but for other reasons.

DR. BRATZLER: Yes, so we actually -- the only reason we use a list of specific organisms is to exclude patients from the measure. In other words, we -- because this measure focuses on empiric selection of antibiotics, if the patient has a documented pathogen, then we actually exclude them from the measure.

And also if they have a documented infection elsewhere that requires treatment, we exclude them from this measure that focuses on only empiric treatment of pneumonia, when you have no pathogen identified.

MEMBER ALMENOFF: I see the exclusion piece. But the inclusion piece is

the part I'm wondering about.

CO-CHAIR GROSSBART: Dale, Steve Grossbart here. Maybe I can help you out. So Peter is referencing the fact that there's septicemia codes and my understanding is that if you have a primary of CAP, you go into this population, or if you have got a primary septicemia and a secondary of CAP, you go into this population.

DR. BRATZLER: Yes, or also a primary respiratory failure and a secondary of CAP, you go in. But again, there has to be documentation by the initial physician, either the admitting, direct admitting physician, or the emergency department physician, that pneumonia was the diagnosis at the time of admission.

So a patient who comes in with respiratory failure, is up on the vent and develops pneumonia three days or four days later, those patients are not included. There has to be a diagnosis of pneumonia up front.

MEMBER ALMENOFF: Yes, but I'm more interested in the diagnosis up front of Pseudomonas pneumonia. Is that, would that be considered a typical community-acquired pneumonia infection, when you get Pseudomonas?

DR. BRATZLER: So Pseudomonas is extremely uncommon as a cause of community-acquired, although there are a few patients, patients with chronic lung disease, who occasionally are diagnosed with Pseudomonas.

But remember, if Pseudomonas was present, if there's a culture positive in that first 24 hours, the patient is actually excluded from the measure, because now we are not talking empiric therapy. We are talking pathogen-directed.

MEMBER ALMENOFF: Okay. Because you know, from the -- from the CMS Hospital Compare experience, because we know that pretty well because we also build a similar model, every year, CMS keeps taking out more diagnoses. And so that's why I'm just kind of

curious, are we going to be doing the same thing again, where we are going to put everything -- every organism on earth on a list, and then because the country starts to groan up or say how could you claim this as a CAP then start to exclude things, I just wondered if we could have, maybe get a little more selective in some of the diagnoses, instead of just taking every single organism on earth and putting it on a big list and calling it CAP. That's just my only concern.

DR. BRATZLER: We do not, we do not use any organism to define the denominator population. No organisms are used to define the denominator. It's an ICD diagnosis of pneumonia, either in the primary or secondary place, with physician documentation of pneumonia at the time of admission.

That's how patients get into the denominator. If they have a specific organism, that actually ends up excluding them from this measure.

CO-CHAIR WEISS: Okay. Is --

MEMBER ALMENOFF: I don't get it but okay. So why are all these on here?

MEMBER JEWELL: I think the confusion is that the list of pathogens to which Peter is referring are above the word exclusions and below the word exclusions on the form, are a list of other criteria.

So it appears as if the list, because I read it the same way, of the pathogens, the diagnosis codes, are the inclusion criteria.

I think that's where the confusion lies.

MEMBER ALMENOFF: So are these all excluded or --

MEMBER JEWELL: That's what I took him to mean. That's -- I think that's the question.

MEMBER ALMENOFF: It's not what it's saying.

MEMBER JEWELL: I get you. I get

you.

DR. BRATZLER: I'm sorry, I just don't happen to have that document in front of me right now. But I can assure you, we don't use pathogens to define the denominator at all.

MEMBER ALMENOFF: Well, we have got to get a clarification again.

CO-CHAIR WEISS: Are we going to need a clarification here?

MEMBER ALMENOFF: Or if it's called something else.

CO-CHAIR GROSSBART: You know, this measure definition has been around for a decade. And I think it's been really aggressively vetted. I mean these ICD-9 codes are inclusion criteria for the pneumonia population, and then there's underlying exclusions for -- and this defines the population for about eight or nine measures.

And I mean these have been out here for a decade, and --

MEMBER ALMENOFF: Yes, they have, but every time, every year, CMS keeps excluding more of them. So if you are very -- if you know a lot of the details of how they do it, and what diagnostic codes are eliminated, so for example, aspiration pneumonia used to be on their list. Now it's not.

A lot of the --

CO-CHAIR GROSSBART: I don't believe so.

MEMBER ALMENOFF: They were. I can -- we build this model every year and we try to be in sync with CMS and --

CO-CHAIR GROSSBART: It hasn't been on the data definition.

MEMBER ALMENOFF: Pseudomonas, I mean, they all get taken off after a while.

CO-CHAIR GROSSBART: It wasn't on, it wasn't on the definitions back in --

CO-CHAIR WEISS: So I'm mindful of how we -- so there's a bit of uncertainty at

the level of where now, just so I make sure the -- and whether or not -- and you are asking, Peter, specifically for --

MEMBER ALMENOFF: Well, he's saying they are not on here. But they are on there. So that's all I need to be clarified.

CO-CHAIR WEISS: Is there some way we can get Dale the document that we are looking at so he can understand what we are talking about? Is --

DR. WINKLER: Dale, do you have our -- are you looking at your computer and can receive email right now?

DR. BRATZLER: I am looking at my computer.

DR. WINKLER: Okay. Can we --

DR. BRATZLER: So do you need me to go to the web -- to your --

DR. WINKLER: Well, either that or we send --

CO-CHAIR WEISS: Can you forward a copy to him?

DR. BRATZLER: I am almost to your site.

DR. WINKLER: Okay.

DR. BRATZLER: But now it says the meeting is not active so I can't get to it.

DR. WINKLER: Okay can you get email?

DR. BRATZLER: I can.

DR. WINKLER: Okay. We are going to see if we can send it to you.

MEMBER GLOMB: Down on page 16 is the only thing that says denominator exclusion details, and it's very brief. It's just cystic fibrosis, in that whole --

DR. BRATZLER: I don't know, the other thing is that you are looking, an ICD-9 diagnosis that includes pathogens. But remember, oftentimes the pathogen is not documented until later during the stay. So we are looking at a long list of ICD-9 diagnoses for pneumonia to find the denominator, but then they also have to have a working

diagnosis for that initial diagnosis of pneumonia when they come in.

If there's a pathogen documented, either through tests like in their antigen test, or a positive culture within 24 hours, they are excluded from the denominator as a performance measure.

CO-CHAIR WEISS: So where are we with in terms of -- Peter, you still seem confused.

MEMBER ALMENOFF: It's fine, don't worry.

CO-CHAIR WEISS: Okay. Then let's go to vote. Validity. Based upon what we have heard so far, high, moderate, low and insufficient.

(Pause for voting)

CO-CHAIR WEISS: And one more time again. Got it.

DR. BRATZLER: Okay, so I did get this table. Table 3.1 you are talking about, on the first page?

CO-CHAIR WEISS: Yes, and as you are looking at --

DR. BRATZLER: Yes, so again, those are only ICD-9 diagnoses that are used to define a denominator population, but again, the measure is only looking at what happens in the first 24 hours of the hospital stay.

So a patient that has documented pneumococcal pneumonia but the blood culture isn't positive until day two or three, they are still in the measure, because initial treatment is empiric.

CO-CHAIR WEISS: Okay. Validity, 10 high, 8 moderate, 1 low and no insufficient. Let's move forward. So we are finally on usability and feasibility. Dave?

Usability and feasibility.

MEMBER RHEW: Yes, again, as Steve has pointed out, this has been around for 10 years or so, and we are currently capturing it through the EHR, through other mechanisms, through paper. It's highly, highly feasible,

and highly reproduceable.

CO-CHAIR WEISS: Any comments on this

MEMBER RHEW: Nothing apart from anything else that we have mentioned already.

CO-CHAIR WEISS: Good. Usability. High, moderate, low, or insufficient. Please vote.

(Pause for voting)

CO-CHAIR WEISS: Okay. If everyone could just please vote again. We are at 18. We are at 19. There we go. Okay. So yes, high, 15, moderate 4, and no low, no insufficient.

And let's go to feasibility as a last one, high, moderate, low, and insufficient.

(Pause for voting)

CO-CHAIR WEISS: We are getting close. Everybody vote again please or vote if you haven't voted.

Done. Good, 19 yes.

It's a good moment. Let's vote on the measure for moving to endorsement. Yes, no, one, two. Yes being one, no being two.

(Pause for voting)

CO-CHAIR WEISS: And while we are doing that, Don, you should be teeing up for the last of this run. Nineteen yes and we are on to Measure 0148, blood cultures performed in the ED prior to initial antibiotic received in the hospital.

Impact. Gap. Evidence.

MEMBER YEALY: So this one we had strong feelings that were not uniformly positive, might be the most charitable way I could frame this.

The concern is, is that the measure as written has no direct link to an outcome, at least not a patient-centered outcome or a particular physician or care provider behavior that could be linked to a patient care outcome, that there were many confounding issues such as the timing of the

two behaviors that are being assessed simultaneously, that could introduce error into the assessment and also produce unintended consequences.

The most specific would be that if, if your goal was to make sure that the blood culture was done before antibiotics, you would separate them in time and space as much as possible to not be, quote, dinged, and in fact, produce an outcome that you didn't want, which is delayed antibiotic therapy, it's not helpful, has to do with that how you measure these two events happen in two different spheres, and even if B followed A, if you are not really careful, it can look like B came before A, and it becomes a problem.

That also then gave some issues about reliability in the timing. It's also not congruent with Measure 0356, which I am sure we will do later, which says if you happen to be sick enough to be in an ICU, you ought to get one in the first 24 hours, but

doesn't make a proclamation about the before and after with antibiotics.

And there's no requirement that these actually be done appropriately, just the timing of things be done first. I guess the driving thing from an impact statement is antibiotics change therapy less than five percent of the -- excuse me, blood cultures change therapy less than five percent of the time.

And the vast majority of that, is in some slight narrowing of antibiotic coverage, not picking something you hadn't already considered.

So for the vast majority of patients, this can't have, this can't have an impact in any way, shape or form. No one is arguing that giving the antibiotic first makes the test better or equal, but in fact the test that isn't useful, it doesn't actually matter which order you do things in.

That's probably the bottom line

here.

CO-CHAIR WEISS: Excellent. Workgroup, thoughts, comments?

DR. BRATZLER: So this is Dale. I don't want to -- I am not going to argue the points about usefulness of blood cultures and again, I highlight that we only have this measure for those patients for which the ED physician elected to do it.

But I would argue that there is usefulness in blood cultures for some patients. You know, if you look at randomly assigned patients, about seven percent will have a pathogen, but if you look at certain populations, when you know, if you take the patient population that's going to be going to the ICU, critically ill patients, patients that have chronic liver disease, hypotensive patients and others, those patients actually do have much higher yield from their blood culture.

So I just would argue that blood

cultures are not useful in the management of pneumonia. Do clinicians use them? Not always, but I think our experts would argue that blood cultures are meaningful for some patients.

CO-CHAIR WEISS: Okay, thank you. I want to make sure that the performance gap, where do we stand with that?

MEMBER YEALY: It actually looks like it's done on a fairly high percentage of cases right now, and again, the get out of jail free card here is if you choose to not draw them, you are off the hook.

And so this -- no one is mistaking that, no one has said that you have to draw blood cultures here. But this becomes an issue that in some ways, you create, if you are very, very efficient, a catch-22, in that if you do these right before the antibiotic therapy, you can give the illusion of having followed it and delivered poor care.

And so as it's written in this

very -- this population, not a much more specific population, it can't possibly deliver the benefit, and likely can only create maladaptive behavior that doesn't benefit any particular patients.

CO-CHAIR WEISS: Tell us what the group really felt on this one. Okay.

MEMBER YEALY: So as a sidelight, this is the criteria that I got the most email before ever joining the --

CO-CHAIR WEISS: And I think we are hearing not just the gap, but evidence in -- response to evidence. So let me ask the group in terms of issue of impact, gap, or evidence. Do we have questions for Don or the workgroup?

MEMBER EDELMAN: Just a clarification.

CO-CHAIR WEISS: Norman, if you could put the mic on.

MEMBER EDELMAN: Just a clarification. So if a blood culture is never

drawn, there's no violation of the standard? Am I right about that?

MEMBER YEALY: Yes.

MEMBER EDELMAN: Okay well, thank you so much.

CO-CHAIR WEISS: So let's vote then, if that was the only question. So high for impact, moderate, low for impact or insufficient.

(Pause for voting)

CO-CHAIR WEISS: Okay let's press again. Oh no, we are all set. This is a squeaker. So high five, four moderate, eight low and two insufficient. That's 10-9. That could mean -- it doesn't pass.

Okay, well, it doesn't pass based upon low impact. From what we were hearing, the low impact is, is that while it does affect patients, it's at the margin, it's a subpopulation of patients that really would be likely affected, and there is evidence to think that when it is effective, it's

effective at reducing the number of antibiotics, not at a change to an antibiotic, that was not currently being administered. That's all anecdotal, but that was what we heard and that was what we acted on.

And that goes back to the measure developer. With that in mind, I wish everyone a 10-minute break. You deserve every minute of those 10 minutes. Thank you all and when you get back, Steve will be helping you usher the next set of random measures. Thank you all.

(Whereupon, the above-entitled matter went off the record at 4:15 p.m. and resumed at 4:25 p.m.)

CO-CHAIR GROSSBART: We are close to an hour behind schedule, and we would like to wrap this up. Ideally we should be open to public comments in 65 minutes, which means we are going to have to move with some speed.

To begin this, to begin this final set of measures that we are going to look at

today, I'd like to ask three measure developers to provide a two-minute summary of the measures that are under consideration and I'd ask that you present all your measures under the COPD section. We'll start with AMA PCPI then we'll move to NCQA and then we'll move to ActiveHealth. And two minutes.

AMA PCPI? Let me start over. So AMA PCPI developer, please give us a quick, two-minute overview of the two measures under consideration, COPD spirometry and COPD inhaled bronchodilator therapy.

AMA, are you on the line?

CO-CHAIR WEISS: It's Dr. Bruce Krieger that we are expecting on the line.

CO-CHAIR GROSSBART: Dr. Bruce Krieger, are you on the line?

(No response)

CO-CHAIR GROSSBART: I tell you what, we'll circle back to you guys, we'll go to NCQA pharmacotherapy management of COPD. Do you have a speaker who can speak to this

measure?

MR. HAMLIN: Yes, this is Ben Hamlin. I am back on the phone. Can you hear me?

CO-CHAIR GROSSBART: Yes, we can. Two minutes.

MR. HAMLIN: Okay. We actually have two measures for COPD. The first one is spirometry testing for a new diagnosis. It's effectively a confirmation of diagnosis testing.

Pharmacotherapy management of COPD exacerbation is an episode-based measure looking to ensure that patients who appear in the ED for an exacerbation are, you know, being prescribed appropriate medications to control their COPD symptoms.

Both measures are administrative-based claims. Both measures have been in HEDIS roughly I believe about five years now each, and they continue to show improvement although there is still room, you know there

is still room -- the gap still exists, excuse me, in the rates that I think, I believe will show up on your sheets.

CO-CHAIR GROSSBART: Are there any questions from the committee to this developer?

(No response)

CO-CHAIR GROSSBART: All right. Is AMA on the phone?

(No response)

CO-CHAIR GROSSBART: Dr. Krieger, are you on the phone?

(No response)

CO-CHAIR GROSSBART: Okay then, we'll move to ActiveHealth. Do we have a spokesperson from ActiveHealth on the phone to discuss their COPD management of poorly controlled COPD?

DR. CHIN: Yes, we are on the line. Can you hear us?

CO-CHAIR GROSSBART: Yes we can.

DR. CHIN: Hi, this is Dr. Lindy

Chin from ActiveHealth management, and we have a team here. Our measure is titled COPD: management of poorly controlled COPD. This measure is looking at the percentage of patients aged 18 years and older who have poorly controlled COPD and are already on a short-acting bronchodilator who are prescribed a long-acting bronchodilator.

Our measure is using claims as well as, where we can, patient self-reported data and health information exchange data as well.

CO-CHAIR GROSSBART: All right. Any questions for the developer from the committee?

(No response)

CO-CHAIR GROSSBART: And finally, Dr. Bruce Krieger.

DR. KRIEGER: Yes.

CO-CHAIR GROSSBART: Yes. We'd like a brief, two-minute overview of the measures that AMA PCPI has submitted.

DR. KRIEGER: Okay. This is Bruce Krieger. I was on the American Medical Association PCPI COPD measures forum which convened about seven years ago, and I was representing the American Thoracic Society. Also present there were multiple other pulmonary societies.

These measures that we are going to discuss were approved by PCPI in 2006. In fact, the measures were previously received and directed to NQF, but they are being reviewed now for maintenance.

The operative COPD, the importance is that, as you all know that COPD is the fourth leading cause of death and that there are recent assessments showing that quality of care delivered to U.S. populations is only -- it's average. Only about 50 percent of COPD patients receive recommended care, but they -- and it was better for exacerbations than for routine care, and nearly 80 percent of COPD patients are undiagnosed, in addition to many

mis-diagnosed patients.

Therefore the two measures that are being presented here, one has to do with diagnosis, which is Measure 0091, which is spirometry evaluation, and the other measure that is being presented is Measure 0102, which is bronchodilator therapy.

CO-CHAIR GROSSBART: Thank you, Dr. Krieger. Is there any questions from the committee at this stage? Dianne?

MEMBER JEWELL: Dr. Krieger, the denominator specifies all patients with a diagnosis of COPD for the spirometry measure. Is that newly diagnosed COPD, all COPD or both?

DR. KRIEGER: It's -- the measure that the PCPI is proposing is not just newly diagnosed COPD. It's to evaluate patients with COPD as well as newly diagnosed, and the reason for that is many patients are labeled COPD without ever having a spirometric diagnosis to confirm that.

MEMBER JEWELL: Thank you.

MEMBER EDELMAN: I don't understand the goal of the spirometry proposal. Is it to capture undiagnosed COPD, or overdiagnosed COPD?

DR. KRIEGER: It's actually designed to capture patients who have a diagnosis of COPD, because the trigger is the patient with COPD with a measurement, both the numerator and the denominator.

MEMBER EDELMAN: I don't understand. If this is intended for quality improvement, it has to correct a mistake that's being made. What mistake are you trying to correct?

DR. KRIEGER: Could you repeat? I did not catch that.

MEMBER EDELMAN: I don't understand. If this is a measure to improve quality of care, it has to improve a mistake, presumably a mis-diagnosis, so you are trying to improve the under-diagnosed COPD, or

correct the over-diagnosed COPD, that is people who have a diagnosis of COPD but don't have it?

DR. KRIEGER: Actually it's both, because the recommendation is that spirometry should be performed in all patients suspected of having COPD.

So it's not -- it will also include patients who do not have the label of COPD but are suspected, and therefore will improve care of both patients.

In addition it will be performed -- it will help diagnose patients with other entities who might have been mislabeled as COPD.

CO-CHAIR GROSSBART: But just a point of clarification, it will only measure those with a diagnosis of COPD?

DR. KRIEGER: No, I'm sorry. I misstated that. It is those suspected of having COPD as well as those who have COPD.

CO-CHAIR GROSSBART: The

denominator states patients with a diagnosis of COPD. So it might prevent the misdiagnosis, but they will fall out of the measure. It's not a bad thing.

DR. KRIEGER: That is correct, but it will also diagnose patients who don't -- who are just suspected, and this will confirm a diagnosis so that appropriate treatment can be rendered.

CO-CHAIR GROSSBART: Okay. Thank you very much for that clarification. So now I'd like to move on to our first measure for consideration, which is 0091, and Dianne will take us through that. Do you want to just give a really quick overview and then we will get into the components?

MEMBER JEWELL: Sure. So I think it's safe to say that the crux of the workgroup's conversation related to this measure really revolved around the questions that we just asked, because the guidelines -- clearly there is evidence of high impact.

Clearly the guidelines, or I should say the guidelines are clear about when spirometry is indicated to diagnose COPD, and when it's not indicated to monitor after treatment.

But the notion that the patients in this denominator have a diagnosis of COPD makes the measure cloudy, I think.

There are performance gaps that we'll go over in a moment. But I -- it's just not as clean with the denominator written as it is. So I would look to my workgroup colleagues to see if there are other things they might add.

(No response)

MEMBER JEWELL: Do you want me to go through the --

CO-CHAIR GROSSBART: Go ahead, from the audience.

MS. AST: May I make a comment?

MEMBER JEWELL: Yes, go ahead, sorry.

MS. AST: Sorry, this is Katherine

Ast from the AMA PCPI. And just a little more clarification on that from one of our co-chairs of the COPD workgroup.

He said that COPD is under-diagnosed and also over-diagnosed for patients who are heavy smokers, so the spirometry evaluation confirms either of these cases.

The management is different for different lung diseases so the spirometry evaluation is needed for confirmation of diagnosis.

I don't know if that helps. But it's for both. So you said, is it under or over. It's both.

MEMBER ALMENOFF: You have a lot of people with a diagnosis of COPD. It's the same thing as the person having a diagnosis of heart disease without an EKG.

So it's a diagnosis but it's really not a diagnosis because they really never validated it with spirometry, which is part of the package.

So they see an x-ray with emphysema and a person who smokes and they give them a diagnosis. That's not a diagnosis. They need to have some definitive testing and that would be spirometry.

Let me just -- I think that's, I think, the issue, that I think those people are bringing up. So I mean, you have a lot of suspected diagnosis of COPD who of course have an x-ray, so they will all -- anybody with a diagnosis, a supposed diagnosis will get screened and if the spirometry is absolutely normal, it probably is not, and then the second phase is that people with diagnoses, supposed diagnoses who have spirometry that validates it, then you make a real diagnosis.

So a lot of people with not real valid diagnoses --

MEMBER JEWELL: So this is probably getting ahead of the order that we normally go, but I think an example of how it could be cleaner is that in the exclusion --

the exclusion criteria are very broad. Documentation of medical reasons, you know, system reasons and so on for why spirometry is not documented.

I would think that already having prior results for spirometry in the record would be a very specific exclusion that should be highlighted so that people don't mistake this measure as I already know that this is an affirmed diagnosis. I have met all the diagnostic criteria, but I am supposed to keep monitoring because this measure says persons with a diagnosis of COPD should have spirometry testing done.

That's my worry about the measure, truly, I mean, just to cut to the chase. It's not clear enough to indicate that what you are not proposing is that this is to monitor people with an affirmed diagnosis, because the guidelines are clear that there is no evidence to support that.

CO-CHAIR GROSSBART: Let's step

through the evaluation process. So Dianne, impact, opportunity and evidence.

MEMBER JEWELL: So, with all that in mind, you know, there's no doubt about the impact. There's more than enough information out there to reflect the incidence and prevalence of COPD and the cost in both quality and literal cost of the disorder, particularly if the diagnosis is missed or not being managed well.

In terms of potential gaps, this measure does have some suggestion of gaps, 45.7 percent of patients reported did not meet the measure, but I offer that with all of the concerns that I expressed a moment ago, because I don't know, really, what's missed here, or what behavior is being captured, I think is a better way to say it.

In terms of the evidence, again, the guidelines are very clear about the indications for which spirometry are most useful, so assuming that those are the

behaviors we are after, I don't think there's any doubt there.

CO-CHAIR GROSSBART: Okay. Any comments from the workgroup?

(No response)

MEMBER JEWELL: I think it's just the two of you. I think Christine -- Christine was the third, I think.

CO-CHAIR GROSSBART: That's right. Good point. Any comments from the larger committee?

MEMBER LARSON: Well, they have a paragraph here that says out of 500 U.S. PCPs, 70 or 69.1 percent agreed that, when COPD is suspected, the diagnosis should be confirmed by spirometry.

So that's sort of like the crux of it, it's primary care, I believe. That's compelling to me.

CO-CHAIR GROSSBART: Let's -- if there's no other comments, let's move to voting. Oh, sorry.

DR. ANTMAN: If I may, just a comment regarding the question about the exceptions specified in the measure, to the earlier question.

As Dr. Krieger pointed out on the phone, these measures were developed a number of years ago, and at that time, the PCPI methodology was to allow for medical patients for system reasons, allowing for clinician judgment, but without providing examples.

So the example that you provided of a specific reason for excluding or accepting a patient for a measure, that is certainly a clarification that we can add, and so we are happy to take that back to our workgroup to consider.

CO-CHAIR GROSSBART: And actually Reva brought up one thing in a quick sidebar. Some clarification is needed in the numerator statement. It's patients with documented spirometry results in the medical record and then the numerator time window, at least once

during the measurement period, and I know the workgroup had questions about potential overuse or unnecessary testing and so on. Did I capture that right? And so if you could address that concern about the time window.

MS. AST: Yes, I'm not sure if it's in your packet or not. In the numerator details, we have numerator instructions which says, look for the most recent documentation of spirometry evaluation results in the medical record. Do not limit the search to the reporting period.

So it's not intended to repeat the spirometry if it has already been done once, ever.

CO-CHAIR GROSSBART: So -- so once a year, you look at your performance but you can look at prior, prior measurements.

MS. AST: Correct.

CO-CHAIR GROSSBART: Okay. Okay so with that, let's get our vote controls out. Importance of the measure and, excuse me, the

impact measure, high is one, medium, moderate is two, low three, four for insufficient evidence.

(Pause for voting)

CO-CHAIR GROSSBART: And so we have 16 high, 2 moderates and zero for the other two categories. Moving to opportunity. Again, one is high, two is moderate, three is low. This is opportunity or performance gap.

(Pause for voting)

CO-CHAIR GROSSBART: Score of 12 for high, 4 for moderate and 2 insufficient evidence. And then our final question, this is a yes/no. Is the evidence sufficient? One yes, two no, three insufficient.

(Pause for voting)

CO-CHAIR GROSSBART: And 16 yes, 2 insufficient. I would like the record to note that the scores are coming up much faster with the new co-chair.

(Laughter)

CO-CHAIR WEISS: What were you

doing when you were not co-chairing?

CO-CHAIR GROSSBART: Okay then, moving on. Reliability and validity.

MEMBER JEWELL: I actually -- could you -- you said something a second ago in response to the numerator time window that I just need clarification on. So I'm looking at that same sentence and it says, look for most recent documentation of spirometry results in the medical record. Do not limit the search to the reporting period.

And you said they wouldn't look at it again if it had been done prior or repeated. But I'm not clear how that's true here. Did I misunderstand what you said?

If I look in three years, and I'm reporting, and I look back and the most recent one was, you know, two years priors, well there could have been 10 before that. I am capturing the most recent ones.

So the notion of continuous monitoring even when it's not indicated could

still occur, right?

MS. AST: It's not intended to have the test repeated, so we can certainly clarify that language if it's still confusing.

MEMBER JEWELL: I would think that that would be one interpretation, that most recent doesn't by itself mean -- it doesn't say only the most recent and no more. If they've got it, then you're done. Continuing to report implies there's more to report, to me. Maybe I'm the only one thinking that way. I'm seeing some heads shake around the group so if I'm the only one, I'll stop perseverating on it, but that's -- okay stop? Got it. All right. Stop. Everybody is more comfortable with it than I am. All right, so -- no, let me just finish.

CO-CHAIR GROSSBART: Reliability and validity.

MEMBER JEWELL: Reliability and validity, yes, I think, with all of that in mind, there were no concerns specifically from

the group that I remember.

CO-CHAIR GROSSBART: Any comments from the rest of the workgroup? Any questions from the larger committee?

(No response)

CO-CHAIR GROSSBART: Okay. So for the question of reliability, one is high, two is moderate, three is low, four is insufficient.

(Pause for voting)

CO-CHAIR GROSSBART: So we have nine high, eight moderate and one low. And then validity. Dianne.

MEMBER JEWELL: Yes, there were no concerns that I remember. I am looking back here at this workgroup list. So I think we are all right.

CO-CHAIR GROSSBART: Although it does show --

MEMBER JEWELL: Oh I'm sorry, I'm looking -- that's because I'm looking at the wrong measure. My apologies. I've got two

things running here. Let me get to the right one.

CO-CHAIR GROSSBART: Yes, the workgroup was not enthusiastic --

MEMBER JEWELL: So the questions, I was looking at the home health measure. So some issues about the validity testing. It's only been tested in one academic medical setting, and there's still -- it was some of the questions -- some of the issues around voting were the questions that we asked about before relative to overuse, so I think that explains why there was a mixed bag.

CO-CHAIR GROSSBART: So again, we'll re-vote on reliability. Validity, rather.

(Pause for voting)

CO-CHAIR GROSSBART: We can always re-vote. Eighteen. Let's all vote one more time. And so we have nine high, seven moderate, one low and one insufficient. And now we'll move on to usability and

feasibility. So usability, Dianne.

MEMBER JEWELL: So this is already a part of the PQRS system as I recall. And I don't know that we have any, any data from the developers per se about how it's performing under those conditions in terms of the public understanding or what have you, but it is in use already.

CO-CHAIR GROSSBART: Any questions or comments from the larger workgroup? Questions by the committee?

(No response)

CO-CHAIR GROSSBART: Well let's move on to voting for usability. This is a one to four range again.

(Pause for voting)

CO-CHAIR GROSSBART: That's our 15 minutes. Okay. So the results were nine high, seven moderate, one low, one insufficient.

And then next, feasibility.

MEMBER JEWELL: Nothing, nothing

concerning leaped out at the group that I recall.

CO-CHAIR GROSSBART: Workgroup, any questions, comments? Larger committee, any questions about this?

(No response)

CO-CHAIR GROSSBART: Then again we will vote on a one to four scale. Did that time out already? We are going to have to vote again.

(Pause for voting)

CO-CHAIR GROSSBART: What's our count up to, 17? Please vote again if you have not. And we had, for feasibility we had 10 high, 6 moderate and no other votes.

MS. WEBER: That's actually eight moderate.

CO-CHAIR GROSSBART: Eight moderate, I'm sorry. And it looks like we had three lows and four insufficient.

(Laughter)

CO-CHAIR GROSSBART: And then the

final yes/no vote on the endorsement, one for yes, two for no.

(Pause for voting)

CO-CHAIR GROSSBART: And the final vote was 17 in favor and 1 opposed. Okay, moving on to our next measure, Dianne, you also have this one, the use of spirometry for the assessment and diagnosis of COPD.

MEMBER JEWELL: Right, so this measure from the NCQA is similar to the measure we just considered, except that it is clear that it is focusing in on the new diagnosis of COPD and the use of spirometry to confirm that diagnosis.

I didn't reference this specifically, but in the prior measure the age range was 18 and I forget what the upper limit was. The initial range here is actually 40 years and older, so that's another distinction.

But really the evidence base is the same in terms of impact and in terms of

support from the guidelines.

The NCQA's data also maps out according to -- in terms of where the true performance gaps lie, maps out commercial, Medicaid and there's one more that I don't have in front of me right this minute, but it's clear that there's a gap based on the data that they have. I want to say it ranges from something like 20 percent to 50 percent.

DR. WINKLER: It's on page 13.

MEMBER JEWELL: Thank you. So really reliability and validity testing is present there, and I think really the workgroup's question perhaps was, revolved around why the cutoff at 40, I think was one of the questions raised. I don't remember which one of us raised that.

It's part of HEDIS so it's in use.

MEMBER GLOMB: Dianne, in that denominator statement, is that just a misprint, the 42?

MEMBER JEWELL: Actually, we might

need some clarification from the NCQA on why they say 42. I think it has something to do with when they capture the data for the person who was 40. But is the developer on the phone or here?

MR. HAMLIN: Yes, this is Ben, I'm here. So two things. First we say 42 in the description because there's a negative diagnosis period, a look-back period to ensure that it's actually a new diagnosis, that this is confirmation of new diagnosis of COPD using spirometry.

The other thing is the reason that we select 40 was for two reasons, one because there's a certain specificity issue with the 18 to 40 group in using spirometry. There's also a concomitant diagnosis of asthma issue, so the amount of noise in the data, we have done a series of analyses based on sort of concomitant diagnosis from you know, 40 through 56, and we have decided that 40 is an appropriate age range and the data is clean

enough and reliable enough at that age for a COPD diagnosis, for us to keep it as our lower limit.

Below that, the noise in the data becomes above our threshold of comfort.

MEMBER GLOMB: Mathematically, though, couldn't you be in the numerator without being in the denominator?

CO-CHAIR GROSSBART: Mathematically can you be in the numerator without being in the denominator?

MEMBER GLOMB: If you're 41 and you have been diagnosed and you have had spirometry, you'd be in the numerator, but you still wouldn't be in the denominator.

You'd be counting someone who is not in your total group.

CO-CHAIR GROSSBART: NCQA, can you clarify that? Can you be in the numerator without being in the --

MR. HAMLIN: I'm sorry. I didn't quite hear the question. It was too quiet.

CO-CHAIR GROSSBART: The question was can you be in the numerator -- because of the age criteria, can you be in the numerator, that is be less than 42 years old, but not in the denominator. Or at least there seems to be some lack of clarity around the numerator and denominator statements.

MR. HAMLIN: No, so you actually have to, we would calculate eligible population first and then do the calculation for the numerative compliance.

So people, for eligibility in the numerator, must first meet the denominator criteria with a diagnosis, but almost must have the clean look-back period with no other diagnosis of COPD in it. So that's why there's an age range, I believe it's two years. So that's why there's the 40 and 42 issue on the age side.

CO-CHAIR GROSSBART: All right. Thank you. So let's step through our assessment unless there's any other comments

from the committee or workgroup. So the first question for us to address is impact.

Did you already do that?

MEMBER JEWELL: Well, there's really no difference in terms of what I presented prior. So --

CO-CHAIR GROSSBART: Okay so let's vote, quickly.

(Laughter)

(Pause for voting)

CO-CHAIR GROSSBART: If you haven't voted, vote again. So the vote is, on impact, 12 high, 5 moderate, 1 insufficient. Let's move on to the question of performance gap opportunity, again a one to four scale. Dianne, do you have anything to add?

MEMBER JEWELL: Just to clarify that they have data on commercial Medicaid and Medicare patients and so -- and there's evidence of a gap, for sure.

CO-CHAIR GROSSBART: Thank you. Any other comments from the committee?

Workgroup?

(No response)

CO-CHAIR GROSSBART: Let's vote.

(Pause for voting)

CO-CHAIR GROSSBART: Here we go, 14 votes for high impact or high opportunity, 4 for moderate impact, none for low or others.

So now we are going to move on to the evidence and this is a yes/no question, one yes, two no, three insufficient. Anything to add?

MEMBER JEWELL: Guidelines are clear.

(Pause for voting)

CO-CHAIR GROSSBART: And it was 18 in favor. Let's move on to reliability and validity. Dianne, any --

MEMBER JEWELL: Yes, hang on one second. Yes, as I mentioned, one of the questions that has already been addressed was the issue of why stop at 40 but we have had that answered, and that was really the

principal thing.

I guess from the validity standpoint, another question was the issue of disparities, because there is evidence of disparities, but the NCQA currently does not feel that that - they could incorporate that into this measure because it would be overly burdensome.

So I don't know what their plans are for the future, but from their own application it appears that they acknowledge that it needs to be addressed somehow.

CO-CHAIR GROSSBART: Developer, do you have a comment?

MR. HAMLIN: Yes, so -- I'm sorry.

CO-CHAIR GROSSBART: Go ahead.

MR. HAMLIN: Okay thank you. Yes, no we are very interested in the disparities issue. Unfortunately right now in our -- we continually retest this issue. In our data we have repeatedly found a great variation in the plans' collection of a standardized, you know,

race, ethnicity, SES data.

And so therefore we are not able to require a reporting out of that information alongside these results. You know, we found a variation from zero to 100 percent. Some plans are actively not collecting the data due to legal reasons. Others are very interested in collecting it in a very standardized fashion.

So we will not require it for the measure until we can actually get a level of consistency that we are comfortable with.

CO-CHAIR GROSSBART: Thank you. Well let's move on to our vote. This is for reliability. A one to four scale again.

(Pause for voting)

CO-CHAIR GROSSBART: There we go. And 12 votes for high and 6 votes for moderate. And now validity. Again, a one to four scale.

(Pause for voting)

CO-CHAIR GROSSBART: There we go,

13 votes for high and 5 votes for moderate. Let's move to our usability and feasibility discussions. Dianne.

MEMBER JEWELL: So as I mentioned, it's already been in use in HEDIS for a period of time. There were questions about the extent to which it is informing quality improvement efforts so that was really the workgroup focus if you will.

PARTICIPANT: What was the answer?

DR. WINKLER: We don't know. In terms of the data reported on page 13, they give you three years' worth of data and for the commercial results, the mean in 2008 was 37.6, in 2010 it was 41.7. So you are seeking gradual improvement over time.

PARTICIPANT: (Off mic)

CO-CHAIR GROSSBART: And I think, you know --

PARTICIPANT: (Off mic)

CO-CHAIR GROSSBART: And that rate of improvement compared to a lot of publicly

reported measures is pretty slow.

MR. HAMLIN: Yes, I didn't hear the previous question -- this is Ben again -- but I think the one issue that we struggle with is trying to ensure that the source of the diagnosis code for COPD is appropriate.

A couple of years ago, I think it was three years ago, we refined that to limit that because we were finding a lot of noise in the data from, you know, COPD showing up from respiratory techs who were going in to do some inpatient procedures, where you know they would sort of write COPD on the chart.

We refined the definition. We do see still that there's a gap for improvement, but again you know the limitations of administrative claims codings for us I think is probably one of the reasons that we are having a hard time seeing large increases in the rates now thinking that the gap has been identified and there's obviously some need to improve.

CO-CHAIR GROSSBART: That, plus no one is doing it.

MR. HAMLIN: You have the reason, yes.

CO-CHAIR GROSSBART: So let's move on to our voting. So, usability. A one to four scale again.

(Pause for voting)

CO-CHAIR GROSSBART: And we have a vote of 7 with a score of high, 10 with a score of moderate and 1 with a vote of low, no insufficient. And then feasibility, Dianne.

MEMBER JEWELL: So yes, the workgroup really didn't have a lot of commentary about this. Let's see. Entered for billing purposes rather than part of the care delivery process, there does not appear to be a strategy to migrating eSpecifications. Some question about whether there are potential problems related to gathering this data, that it wasn't clear from the application. But there was nothing that leapt

out from our conversations.

CO-CHAIR GROSSBART: Any comments from the workgroup or the larger committee?

(No response)

CO-CHAIR GROSSBART: Okay let's move on to voting for feasibility.

(Pause for voting)

CO-CHAIR GROSSBART: And the results were 12 high and 6 moderate. And now we get our final yes/no vote, one for yes, two for no.

(Pause for voting)

CO-CHAIR GROSSBART: And it's unanimous, 18 votes in favor. Our next measure is Measure 0102, inhaled bronchial dilator therapy, and Dr. Edelman, you are up.

MEMBER EDELMAN: So I apologize to my workgroup because I thought this was really simple until I reread it. So I am going to read the numerator and denominator because that's where my questions are.

So the numerator is patients

prescribed an inhaled bronchodilator at least once, and the denominator is 18 years old plus a diagnosis of COPD plus an FEV1/FVC ratio of less than 70 percent, plus they have symptoms and the timeframe is 12 months.

The impact we needn't discuss. The impact of COPD is very high. The improvement is where I have a little rethinking.

So if you look at each of the individual elements of the denominator and then you look at the literature that is cited, there is a good amount of evidence that bronchodilators improve function and there is a good amount of evidence that lots of people who meet the individual criteria are not getting bronchodilators.

So that addresses both the impact and the opportunity for improvement. What I couldn't find is evidence that taking the denominator as a whole, that is diagnosis of COPD plus FEV1/FVC ratio less than 70 percent,

plus have symptoms, I couldn't find people who meet those criteria are not getting bronchodilators, and I suspect that those data are not available.

So what appeared initially to show a huge gap and a lot of opportunity for improvement is now unclear to me.

The rest is pretty straightforward. The discussion of disparities is good. The quantity of the data is from a review of nine studies, not all were significant.

The quality of data is buttressed by the report of the ACP, ACCP, ATS, ERS, strong recommendations, and there's a lot of good stuff about reliability and validity.

So you know, we'll go through the individual elements. In general I'm favorable except that I would like to ask the proposer about my question about the gap.

CO-CHAIR GROSSBART: Okay. Can we save that for that section in the conversation

or do you want to hear the answer now?

MEMBER EDELMAN: Well --

CO-CHAIR GROSSBART: Either way.

MEMBER EDELMAN: Why don't we do it now?

CO-CHAIR GROSSBART: Okay, so -

MEMBER EDELMAN: The developer.

CO-CHAIR GROSSBART: The developer, AMA, a question about the performance gap.

MS. AST: I'd like to ask if Dr. Bruce Krieger is still on the phone, if he has any comments about what Dr. Edelman brought up.

CO-CHAIR GROSSBART: Dr. Krieger did you hear the question?

DR. KRIEGER: I heard most of the question, having to do with the denominator and the -- including patients with COPD whose CT barometric definition, which is an FEV1/FVC ratio of less than 70 percent, and had symptoms.

That basically is the starting point in all the algorithms, be it from the global initiatives of obstructive lung disease, the goal ATS, COPD and Canadian, for giving treatment with -- you're giving treatment with COPD and that first line of treatment is a bronchodilator.

MEMBER EDELMAN: No, I -- I'm sorry. Go ahead. I understand.

DR. KRIEGER: I may have missed the question. Oh, as far as the denominator, that basically is the population that should be treated with bronchodilators. Not everyone with COPD needs bronchodilators.

MEMBER EDELMAN: I understand all that and agree with it. But what is the evidence that a person, a group of people, who meet all three criteria as written in the denominator, that is have symptoms and have abnormal spirometry and have a diagnosis of COPD, what is the evidence that a significant number of those people are not getting

bronchodilators?

DR. KRIEGER: There's a study in quality of obstructive lung disease care for adults in the United States published and checked in 2006, showing that COPD patients -- only 58 percent of COPD patients received appropriate care, based on these guidelines.

MEMBER EDELMAN: Was COPD defined by all three criteria in that study?

DR. KRIEGER: Yes.

MEMBER EDELMAN: Okay.

DR. KRIEGER: I think it was Mularski was the lead author, that was the criteria for diagnosing COPD.

MEMBER EDELMAN: All right, that's fine. That satisfies my question.

CO-CHAIR GROSSBART: Thank you. So let's step through our voting elements. We start off with impact.

MEMBER EDELMAN: I think the impact is high. I don't think there's a need to discuss it very much.

CO-CHAIR GROSSBART: Okay, any questions or comments from the committee or workgroup?

(No response)

CO-CHAIR GROSSBART: Then can we initiate the voting, Jessica.

(Pause for voting)

CO-CHAIR GROSSBART: So in terms of impact, the vote is 16 with a score of high and 1 with a score of moderate, no lows, no insufficients. And then performance gap?

MEMBER EDELMAN: With the developer's clarification, I think the performance gap is high.

CO-CHAIR GROSSBART: Any questions or comments from the workgroup or committee?

(No response)

CO-CHAIR GROSSBART: Then we'll vote. Again a one to four scale.

(Pause for voting)

CO-CHAIR GROSSBART: Why don't we try voting one more time, just to get the last

one in there. There we go, 13 votes for high, 4 for moderate. Moving on to our next area, evidence. Did I just skip one? No. Evidence, correct. And this is a yes/no question.

(Pause for voting)

CO-CHAIR GROSSBART: And the score was 17 yes that the evidence was sufficient and 1 no.

Move on to the questions of reliability and validity.

MEMBER EDELMAN: There was a good discussion of reliability which they deem to be moderate.

CO-CHAIR GROSSBART: Let's vote on that. First of all, any questions or comments?

(No response)

CO-CHAIR GROSSBART: Okay let's vote.

(Pause for voting)

CO-CHAIR GROSSBART: And a score

of 7 votes for high and 11 votes for moderate, no other votes cast. And then the validity question again, a one to four scale. Any comments?

MEMBER EDELMAN: No, I think it rolls up to a high validity.

CO-CHAIR GROSSBART: Okay.

(Pause for voting)

CO-CHAIR GROSSBART: I can't see that far. What are we up to? Two more votes. Let's everyone vote one more time. There we go. So the vote was 14 high, 4 moderate.

And then usability, any comments about usability?

MEMBER EDELMAN: I think it's straightforward.

CO-CHAIR GROSSBART: All right. Any comments from the larger workgroup or committee?

(No response)

CO-CHAIR GROSSBART: Okay, well, hearing none, let's vote.

(Pause for voting)

CO-CHAIR GROSSBART: And the results for usability are 16 high and 3 -- 15 high and 3 moderate. And then feasibility. Again any comments?

MEMBER EDELMAN: It is feasible.

CO-CHAIR GROSSBART: And any comments from the workgroup or the committee?

(No response)

CO-CHAIR GROSSBART: All right. Let us vote.

(Pause for voting)

CO-CHAIR GROSSBART: And the results are 14 high and 4 moderate, and now -- for feasibility. And now for overall vote, yes/no question, one yes, two no for endorsement.

(Pause for voting)

CO-CHAIR GROSSBART: And the final vote was unanimous, 18 votes. All right. It looks like I am up for Measure 0549, pharmacotherapy management of COPD

exacerbations.

Let me just get over there. I mean the main points of discussion from the workgroup was that the evidence was strong, essentially the same evidence that we have already discussed for the other measures, and some concerns about the reliability and validity testing, concerns that there hasn't been a trend over time available, concerns about the fact that it's claims-based data and that no eSpecifications were offered, and the last point we'll discuss tomorrow on related and competing measures.

So this measure is based on patients who are -- inpatient or ED visits, and who are dispensed a corticosteroid within 14 days of an event and a bronchodilator within 30 days of event.

And again, similar questions about you know, similar age group as the other NCQA measure. Those are the high points that I had. First I'd ask the workgroup if there's

any comments they'd like to add, as well as to open this up to the larger committee.

MEMBER YEALY: One question, a clarification.

CO-CHAIR GROSSBART: Yes. Yes.

MEMBER YEALY: How are we determining from the numerator the dispense of the medication, particularly as institutions go to handing, you know, the first set of inhalers out, if you are using claims-based data it would be very easy to miss that quality initiative and rebrand it something else. I'm just -- any clarification on how it's being extracted?

CO-CHAIR GROSSBART: In my reading of the specifications, it is claim-based so I'll ask the measure developer to comment on that issue.

MR. HAMLIN: Yes, this is an administrative, claims-based measure only. So it's the health plan collecting the data both from the hospital setting and from the

provider setting.

CO-CHAIR GROSSBART: So if the patient receives medications directly from the provider, it will be a false negative?

MR. HAMLIN: It's -- no, the dispensed prescriptions will show up. It's a little unsure, numerator compliance, if they are actually in fact dispensed them for provider prescription following a discharge from the ED or from an inpatient setting.

CO-CHAIR GROSSBART: I'm still not sure what it means, to be honest with you.

MEMBER STEMPLE: And I'm confused as the patient already has the medications, so what's the false -- what's the false -- is it a false positive because they already have the meds, so what -- I don't see how this measure has much validity at all.

MR. HAMLIN: So, if the patient is actively on a medication already, that actually does count towards numerator compliance and that is actually found to be in

the administrative claims record, so that will count.

MEMBER STEMPLE: So there's a look back for a pharmacy fill, or what's the look back to determine -- can you define that a little bit better, how is that authenticated?

MR. HAMLIN: Well, we get an annual claims dump for the calculation of the measure, so we are looking at all you know, claims processed between January 1st and December 31st and we usually have about a three or four month period before claims are due to us, so we allow the claims to run out in that regard as well.

MEMBER STEMPLE: So does it look back for any script in the previous year or 60 days or 90 days or what's the specificity of the look back to see if they would probably have, already have access to the products that you are looking for?

MR. HAMLIN: If they have an active prescription, so if there was a

prescription dispensed 30 days before, I believe that would be counted as active. I am not sure about longer times, given the medications that are -- you know, the corticosteroid medication, prescribing that -- talked to our vendor, talked to the steroid vendor folks about what the timeframe is.

MEMBER STEMPLE: I think there is some recommendations in sub-guidelines that members are just sort of stockpiled with these as a standing, to sort of supplement if them feel an exacerbation coming on. So is there any data to show how many people just have a ready stockpile so your look back of 30 days will not be valid.

MR. HAMLIN: Right, I don't know what the -- I don't think there's an actual look back period. I think it's just if there's, it's like I said, if there's an active medication during the exacerbation that counts towards numerator, I'd have to look and see what the actual attribution, what the

attributed attribution of a claim for medication dispensed towards the event would be, and I don't have that information. I will have to get that for you.

CO-CHAIR GROSSBART: Any other questions from the committee? Yes Peter.

MEMBER ALMENOFF: Under the systemic steroids, is that inhaled and oral or just oral? Just want a clarification for the -- what steroids are, systemic steroids.

MR. HAMLIN: It's all med classes that are, that are listed on the table PCEC which I am looking for the page number right now for you.

MEMBER ALMENOFF: I don't have the table so I don't know the answer. It's oral and inhaled? Or just various? Okay.

MR. HAMLIN: I believe it's pages 8 and 9 list all the medication classes.

DR. WINKLER: Go back up Katie, to the meds.

MR. HAMLIN: They're specifically

on page 8, for the numerator.

MEMBER ALMENOFF: Okay, so that's both. Okay. Good, thank you. It's got inhaled and oral.

MEMBER YEALY: I would like to know how it would be handled if after leaving the emergency department I gave you your -- either your inhaler or your four days of oral steroids. I dispensed, you know, you got that as part of your ED visit, and then returned to your normal regimen. It's not entirely clear to me that an administrative claims-based would identify that, yet it would be -- in some ways it's actually the most efficient care. I'm certain you have the medicines that you need and that's my only concern about the validity of this. How would you handle that or can you handle that or have you considered that?

CO-CHAIR GROSSBART: So, a question to the developer. You have a patient who receives medications directly in the ED,

not a prescription, not something they have to fill, they walk out with it in their hands. How does your measure account for that? Or can it account for that?

MR. HAMLIN: Yes, if the medication is administered, you know, is dispensed to a person in the ED, that will be captured in the admin claims, because it will be -- it will show up, and therefore it will still be compliant.

MEMBER YEALY: How will -- only if you are charged for it?

MR. HAMLIN: Most of the ED, you know, tend to show up, you know, as CPT under procedural administration, but if they actually get a prescription in the ED, that will actually show up on the admin claims, so they are linked to the CPT codes which are how the measures are reported.

CO-CHAIR GROSSBART: What if they walk out with the actual medications but no prescription?

MR. HAMLIN: Well it's tracked to the medication dispensation which actually shows up in the pharmacy claims, so it's not prescription-based. It's a dispensed-based measure.

MEMBER YEALY: I' remain skeptical, but, I mean, because this cost me a buck to dispense and I'm not sure it hits a charge line. Yes.

CO-CHAIR GROSSBART: It's kind of like aspirin at discharge or aspirin on admission rather. What, they charge you for that? No, no one charges for that.

Anyway, so -- so, well let's move on to our assessment of it, unless there's any questions let's move on to our assessment, beginning with impact, the workgroup thought this was a high impact, largely because COPD is such a high impact disease.

The only question that we had was that there was limited evidence presented that there was underutilization of

pharmacotherapy management.

I'm going to walk over -- so let's -- that's impact but again the committee rated the impact high. Any other comments from the workgroup or questions from the larger committee?

(No response)

CO-CHAIR GROSSBART: Well then let's vote on this first item.

(Pause for voting)

CO-CHAIR GROSSBART: Here we go. So we have 15 with a score of high and 3 with a score of moderate. No other votes. Moving on to performance gap, the committee did not see the significant performance gap so if you look in the measure information --

DR. WINKLER: They're now on page 14.

CO-CHAIR GROSSBART: So you have performance depending on what type of payers, commercial results, in a 70 percent range, variation from 60 to 78 percent for some of

these in 2009, rates higher for bronchodilator, less for corticosteroids, similar results for the Medicaid population, and again, there was the question about is this seriously underutilized, concerns about the definition and so on that we have already discussed.

Any comments from the workgroup, or questions from the committee around the performance gap?

(No response)

CO-CHAIR GROSSBART: So let's vote on performance gap, one to four scale again.

(Pause for voting)

CO-CHAIR GROSSBART: And 2 votes for high, 13 votes for moderate, 2 for low and 1 for insufficient data. And the final area under this, the importance to the measure and report is the quality of the evidence, and this is a simple yes/no.

The committee itself found that the evidence for this measure was rated about

moderate on most categories, as you can see on the report. Any questions or comments from the rest of the workgroup or from the committee itself?

(No response)

CO-CHAIR GROSSBART: We are all tired aren't we. So it's a yes/no, is the evidence sufficient.

(Pause for voting)

CO-CHAIR GROSSBART: So the score, voting was 15 yes, 1 no, 2 insufficient. Going to move to the reliability and validity questions. In terms of reliability, this is a mix of administrative and clinical data.

We have raised concerns about the -- some issues around is the data capable of accurately capturing all the availability of medication and some concerns about preexisting prescriptions and as well as dispensing through the ED, and some questions about the ability to, in terms of validity, the ability to only focus in on primary diagnoses of COPD

and patients with a secondary of COPD would be ignored.

Example given was respiratory failure with a secondary of COPD. So in terms of -- and again the committee rated reliability leading towards the medium side.

So are there any comments from the workgroup? Dianne.

MEMBER JEWELL: So I guess a question. So dispensing a sample is not the same as dispensing a prescription? I'm asking.

CO-CHAIR GROSSBART: There's some questioning of that among the committee.

MEMBER JEWELL: Okay, and I guess I thought I heard the measure developer say, so maybe I just need clarification again, that a prior prescription before the exacerbation would count as meeting this measure. Did I hear that correctly?

MR. HAMLIN: Yes that's correct. So if the patient is on active medication they

will count towards the numerator. The measure intent is to ensure that patients who have an exacerbation are on the appropriate medications to theoretically prevent these exacerbations and so we do count the ones who are actively taking the meds.

CO-CHAIR WEISS: And did we hear that right? It's a 30-day look back but not a 90-day? I'm just thinking about pharmacy benefit managers may dispense like three months' worth of this stuff, and yes.

MR. HAMLIN: I don't have the exact number of days that would count. I'd have to look that up and I don't have that information accessible right now. I sent in a request but unfortunately I don't have that easily accessible to me.

CO-CHAIR GROSSBART: Any other questions from the committee?

(No response)

CO-CHAIR GROSSBART: Let's move on to our voting. So, reliability. One to four

scale.

(Pause for voting)

CO-CHAIR GROSSBART: And we had 1 vote for high, 11 for moderate, 5 for low and 1 for insufficient.

And then moving on to the validity. Again, the committee found this as -- scored this in a moderate range. Any other comments from the committee, or the workgroup, or the committee? Then let's vote. Okay, go ahead.

CO-CHAIR WEISS: A question I had, as part of the validity, did they raise these questions about the look back period and whether or not if someone actually had a recent dispensing beyond 30 days, was that part of the discussion of the workgroup?

CO-CHAIR GROSSBART: No it was not.

CO-CHAIR WEISS: Okay.

CO-CHAIR GROSSBART: So if your point is, we had -- we rated it moderate

before these additional questions up.

CO-CHAIR WEISS: I'm just -- it would be helpful at least to me to know that information, because it's just -- it's going to be some level of mis-classification, the question is how much.

CO-CHAIR GROSSBART: Any other comments or questions?

(No response)

CO-CHAIR GROSSBART: Let's move forward with our voting. So, validity of the measure.

(Pause for voting)

CO-CHAIR GROSSBART: Seven moderate, eight low, two insufficient. That stops this.

DR. WINKLER: Yes, that vote of seven moderate, eight low, two insufficient, that stops this measure. It doesn't pass scientific acceptability.

CO-CHAIR GROSSBART: And I think the -- to sum up the committee's deliberation,

there was some lack of clarity around whether or not patients would be actually getting the appropriate therapy but not being counted as having received that therapy by the measure design.

All right. We have one final measure to go. We have almost made up our lost time. This will be number 1825, a new measure, COPD management of poorly controlled COPD, ActiveHealth.

Norm you are up for this one.

MEMBER EDELMAN: Oh, I love going last. This is not a bronchodilator measure, but to me it's more interesting and better focused. So, the numerator, patients under 18 -- over 18 with poorly controlled COPD who are taking a long acting bronchodilator; denominator patients over 18 with poorly controlled COPD who are taking a short acting bronchodilator; and poorly controlled COPD is several refills of the short acting bronchodilator, a diagnosis of acute

exacerbation of COPD, or refills of systemic steroids.

The impact I think is high, as all these therapeutic issues are in COPD. In this case, I think the gap is well documented and quite clear. It gets much better when it gets specific for a long acting bronchodilator.

I think the quality of evidence is good. Quantity is good. Eight meta-analyses, 42 studies.

Quality -- quality is less good. The developer gets lost in a long discussion comparing long acting bronchodilators which is really not to the point, and I think that turned off one of the members of our subgroup.

But I think the quality of the evidence is good. We'll discuss reliability and validity later. Oh, I have no questions for the developer.

CO-CHAIR GROSSBART: Workgroup, any comments or additional -- and questions from the committee before we go into our

voting phase?

(No response)

CO-CHAIR GROSSBART: All right so let's start with our importance of the measure, so impact. Don't start voting yet.

MEMBER EDELMAN: I think the impact is high. There is a big gap and there's good evidence that long acting bronchodilators reduce exacerbation rates of COPD.

CO-CHAIR GROSSBART: Okay, let's vote, one to four range again.

(Pause for voting)

CO-CHAIR GROSSBART: And the vote was 17 with a score of high. And opportunity for improvement, or performance gap.

MEMBER EDELMAN: As I pointed out, I think that's high.

CO-CHAIR GROSSBART: Questions or comments?

(No response)

CO-CHAIR GROSSBART: Let's vote.

(Pause for voting)

CO-CHAIR GROSSBART: And the vote was 16 high, 2 moderate. Moving on to the evidence. Is the evidence sufficient, yes/no question. Any comments Norm?

(No response)

CO-CHAIR GROSSBART: Okay, let's vote.

(Pause for voting)

CO-CHAIR GROSSBART: And the final vote is 14 yes, 4 no. Let's move on to reliability and validity.

MEMBER EDELMAN: There's a good analysis of reliability, which comes out moderate.

CO-CHAIR GROSSBART: Any questions or comments from the committee, workgroup?

(No response)

CO-CHAIR GROSSBART: Let's vote.

(Pause for voting)

CO-CHAIR GROSSBART: What are we up to there? Let's vote again. And the 3 votes for high and 15 votes for moderate, and

then validity. Again, any comments Norm?

MEMBER EDELMAN: I think it's highly valid.

CO-CHAIR WEISS: This is just a question of age range and I think I may have missed this question on a prior COPD measure, but if you got a 20 year old who was not succeeding at this measure with their COPD, what would you think as a pulmonologist?

MEMBER EDELMAN: Ooh, you are going to ask me a pulmonology question. If I had a 20 year old whose diagnosis is COPD, I would worry about my diagnosis of COPD.

CO-CHAIR WEISS: How about a 28 year old with this process?

MEMBER EDELMAN: I, look, you are raising the question of the interface between COPD and asthma.

CO-CHAIR WEISS: Yes, I'm just wondering --

MEMBER EDELMAN: I mean that's a huge question and that's a question not only

at the lower age range. It's a question at the higher age range.

CO-CHAIR WEISS: So it is a question --

MEMBER EDELMAN: So I think it's a question that runs throughout the age range, and you know, all of these criteria are exceedingly simplistic to a pulmonologist and hopefully only apply to primary care physicians.

CO-CHAIR WEISS: The reason I ask is because we have tomorrow an issue of harmonization of ages, and --

MEMBER EDELMAN: I don't think my -- my answer to your question is I don't think playing with the age profile is going to get you out of the very real problem of distinguishing between asthma and COPD.

CO-CHAIR GROSSBART: All right. Let's move on to our vote on validity. One through four scale again.

(Pause for voting)

CO-CHAIR GROSSBART: And the validity score came out 5 high, 12 moderate, 1 low. Next area is usability. Again there are -- Norm, do you have any comments?

MEMBER EDELMAN: I think it's an understandable and usable metric.

CO-CHAIR GROSSBART: Any other questions or comments from the workgroup or the committee?

(No response)

CO-CHAIR GROSSBART: Okay with that, let's vote. It's a one through four scale.

(Pause for voting)

CO-CHAIR GROSSBART: One more to go, it looks like. There we go. And the score was, the vote was seven high, nine moderate, two low. And then feasibility.

MEMBER EDELMAN: It's easily measured.

CO-CHAIR GROSSBART: So any questions or comments about the feasibility?

If not let's vote. One to four scale.

(Pause for voting)

CO-CHAIR GROSSBART: What are we up to? About 17, 14? Let's everyone vote again. The transcript is going to really be interesting to read.

What are we up to now? Fifteen. No, we had 18 on the last vote, didn't we? Everyone vote one more time. All right. Time is up. We'll see how the results come. If it's close we'll revote.

So, 11 -- we got them all -- 11 high, 2 moderate. The counter could be off.

MEMBER EDELMAN: See the counter is off. That's 18.

CO-CHAIR GROSSBART: And then finally the yes/no endorsement vote.

(Pause for voting)

CO-CHAIR GROSSBART: Eighteen in favor. Yes.

So again, we are 20 minutes behind schedule, but -- and we still have a formal

15-minute session for NQF member and public comments. So I open this -- I would ask any members of the public or -- to comment, if they choose to.

DR. WINKLER: Operator is there anyone on the phone want to make a comment? Operator, are you there?

OPERATOR: Yes ma'am.

DR. WINKLER: Oh good. Does anybody want to make a comment?

OPERATOR: There is no public audience on the phone.

DR. WINKLER: Thank you. All right. Thank you all. You have done an arduous bit of work today. We are at actually not that far off schedule. We were to adjourn four minutes ago, according to the agenda.

So you have all done a fantastic job. However, we still have considerable work to do tomorrow. The agenda, we have 13 more measures tomorrow. A lot of these are outcome measures.

We also need to have a discussion of related and competing measures and now that you have done the first pass review of the process measures, we can take a look to see what's left and see where the issues around competing and harmonization are.

We also, toward the end of the day, if we get through all the measures before everybody has to leave, we do want to have a conversation about gaps.

We see the measures that are here but the question is, what are the measures that should be? You know, what would we like. We have had some input from ACCP on a couple of documents on critical care and pulmonary conditions that we gave to you and are on SharePoint for you to review about gaps in these topic areas so hopefully we will have just a little bit of time.

If, when you should come in in the morning, you let Katie, Jessica or myself know at what point you are planning on leaving so

we can get a sense.

We are hoping to have a critical mass of you all at least until about 3 o'clock, but we do know people will be racing to the airport to catch flights.

Does anybody have any questions or comments at this stage? I'll step out of the way to avoid the rush towards the door. Question? Comment? But again, thank you all very much. It's been a long day. You have been terrific. We appreciate your patience and your cooperation in going through this. Have a nice evening and we'll see you tomorrow.

(Whereupon, at 5:52 p.m., the proceedings in the foregoing matter adjourned for the day.)