NATIONAL QUALITY FORUM

+ + + + +

EYE CARE, EAR, NOSE AND THROAT CONDITIONS (EENT) STANDING COMMITTEE

+ + + + +

WEDNESDAY
JUNE 3, 2015

+ + + + +

The Committee met at the National Quality Forum, 9th Floor Conference Room, 1030 15th Street, N.W., Washington, D.C., at 9:11 a.m., Daniel Merenstein and Kathleen Yaremchuk, Co-Chairs, presiding.

PRESENT:
DANIEL MERENSTEIN, MD, Co-Chair
KATHLEEN YAREMCHUK, MD, MSA, Co-Chair
TAMALA BRADHAM, PhD, CCC-A, Vanderbilt University Medical Center
MATTHEW CARNAHAN, MD, MS, The Permanente Medical Group
SCOTT FRIEDMAN, MD, Florida Retina Consultants
SETH GOLDBERG, MD, Aetna Insurance Company
JUDITH LYNCH, NP, American Association of Nurse Practitioners *
RICHARD MADONNA, OD, SUNY College of Optometry
VAISHALI PATEL, PharmD, MS, US Health Outcomes, Allergan Inc. *
TODD RAMBASEK, MD, ENT & Allergy Health Services
ANDREW SCHACHAT, MD, Cole Eye Institute, Cleveland Clinic
JOSHUA STEIN, MD, MS, University of Michigan
MICHAEL STEWART, MD, MPH, Weill Cornell Medical College
STEVEN STRODE, MD, MEd, MPH, FAAEP, AR Disability Determination Services

JACQUELYN YOUDE, AuD, CCC-A, Healthcare Performance Partners

NQF STAFF:
ANN HAMMERSMITH, JD, General Counsel
MARCIA WILSON, PhD, Senior Vice President, Quality Measurement
SHACONNA GORHAM, MS, PMP, Senior Project Manager
VY LUONG, MPH, Project Manager
KAITLYNN ROBINSON-ECTOR, MPH, Project Manager
REVA WINKLER, MD, MPH, Senior Director

ALSO PRESENT:
MARY BARTON
ELVIA CHAVARRIA
KENDRA HANLEY
BENJAMIN HAMLIN
JOY JIN
TONI KAYE
FLORA LUM
SAM TIERNEY
RICHARD ROSENFELD
WILLIAM RICH

*** present by teleconference**

TABLE OF CONTENTS

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

9:11 a.m.

DR. WINKLER: Good morning and thank you for joining us. I'm Reva Winkler, I'm a Senior Director here at NQF.

Thank you all for joining us for our EENT measures evaluation and being members of our Standing Committee.

NQF has moved to standing committees, we'll draw choices between two and three year terms for you so that we have a committee available to discussion this topic area when issues come up and when we need to review measures. So, we do appreciate the time that you're volunteering and spending.

To kick off our introductions, I want to introduce our co-chairs, Dan Merenstein and Kathy Yaremchuk. And so, we'll be turning the meeting over to them when we get going a little bit later.

But, right now, I'd like to introduce NQF's General Counsel, Ann Hammersmith, and she

will help us do introductions and disclosures of interest.

MS. HAMMERSMITH: Good morning everyone. As Reva, said, we'll combine the introductions with the conflicts disclosure because it saves a little time. It's a little bit easier.

If you recall, we sent you a rather long form before you were named to the committee where we asked you about your professional activities in some detail.

What we'd like to do now is have you go around the table and declare anything that you wish to declare that you think is relevant to your service before the committee.

Please do not summarize your resume. You don't even need to summarize the form.

The idea is for you, in the spirit of openness, to disclose any activity that you've been engaged in in the last five years that may be relevant to the subject matter before the committee, but only if it's relevant to the

subject matter before the committee.

Before I have you go around the room, I want to remind you that you serve as an individual, as an expert. You're not here representing your employer. You're not here representing anyone who may have nominated you for service on the committee, purely as an individual expert.

Just because you disclose something doesn't mean you have a conflict. You're simply stating things to your fellow committee members and to the public that you think are relevant and important to the work that will be done here today.

We are particularly interested in any research that you've done that's relevant to the topic today, any grants you've received or speaking engagements, but only if it's relevant to the subject matter before the committee.

So, with that, I'll start with the chairs. If all of you could introduce yourselves, tell us who you're with and if you

have anything to disclose.

CO-CHAIR YAREMCHUK: I'm Kathy Yaremchuk, Chair ENT, Henry Ford Health system and I don't have anything to disclose.

CO-CHAIR MERENSTEIN: I'm Dan Merenstein, family physician from Georgetown University.

Part of my research is on sinusitis and I do work on a committee with Pew and CDC on decreasing antibiotic usage and it's primarily ENT issues.

MEMBER YOUDE: Hi, everybody. My name's Jackie Youde. I have nothing to disclose.

I am an Audiologist and I work as a consultant at Lean Six Sigma with Health Care Performance Partners, a MedAssets Company.

MEMBER SCHACHAT: I'm Andy Schachat. I'm a retina specialist at the Cleveland Clinic and these instructions are much broader than the instructions I read.

You're asking about research and grants and I've participated in 10 or 20 research

projects and probably written 200 articles related to the field and I've edited a journal and edited a textbook.

So, I'm just -- I can't begin to list them for you, although I've shared the CV.

MEMBER GOLDBERG: I'm Seth Goldberg, Medical Director from Aetna Insurance and I have nothing to disclose.

MEMBER BRADHAM: My name is Tammy Bradham and I am at Vanderbilt University. I am an NCHAM consultant, National Center for Hearing Assessment and Management.

I have published about, not 200, take a zero, 20 articles in the area of early intervention identification and hearing loss. I do have a current grant right now that's looking at spoken language development in children newly identified with hearing loss.

MEMBER STEWART: I'm Micky Stewart. I'm an otolaryngologist at Weill Cornell Medical College in New York. And I don't have any conflicts.

I am on the Board of -- I'm elected Director of the Board of Directors of the American Academy which is the sponsor or the steward of several of the measures but I've never been actually on a -- I've never developed a measure. I haven't been on the committee. I haven't been an author on any of the guidelines. I'm just on the Board of the Academy.

MEMBER STEIN: My name's John Stein. I'm at the University of Michigan. I'm an ophthalmologist and glaucoma specialist there. I'm on the American Glaucoma Society Board of Directors.

And my research, I'm a health services researcher. I do a lot of large health care claims, database analyses, some of which are relevant to some of the measures.

MEMBER RAMBASEK: My name is Todd Rambasek, allergist with ENT & Allergy Health Services and I have no relevant research or grant interests.

MEMBER MADONNA: Good morning. I'm

Rich Madonna. I'm the Chairman of the Department of Clinical Education at the SUNY College of Optometry.

I do consult for or on advisory boards for or speak for drug companies and instrument companies. However, I don't believe that they will have any direct relevance here.

MEMBER FRIEDMAN: Scott Friedman, ophthalmologist, Lakeland, Florida. I have no significant financial disclosures.

MEMBER STRODE: Steven Strode. I'm a family physician from Little Rock. I did participate with the American Academy of Otolaryngology in developing a clinical measure on hoarseness and then with the AMA on sinusitis. I think both were about four or five years ago.

MEMBER CARNAHAN: Matt Carnahan, an ophthalmologist, the Permanente Medical Group in Northern California. I have nothing to disclose.

MS. HAMMERSMITH: And I understand there are some people on the phone.

Judith Lynch?

MEMBER LYNCH: I'm Judith Lynch from Milford, Connecticut. I'm a Nurse Practitioner in otolaryngology and allergy.

I've written several articles on different issues in ENT and allergy.

MS. HAMMERSMITH: Okay, thank you.

Vaishali Patel?

MEMBER PATEL: I'm Vaishali Patel. I'm a pharmacist in the Health Outcomes researcher by training and much of my research is in ophthalmology health outcomes research. And I currently work at Allergan.

MS. HAMMERSMITH: Okay, thank you.

Before I leave you, I just want to remind you that we rely on you as committee members to help us successfully deal with any bias or conflicts of interest.

So, if you're in the meeting and you think that one of your fellow committee members has a conflict or if you have a conflict or if you think someone is behaving in a biased manner, please do speak up. Please don't sit there and

then tell us six months later, well, you know, I thought so and so had a really huge conflict and we didn't know about it. We want to deal with it in real time.

If you want to say something about you think you have a conflict, you think someone else does or is acting in a biased manner, you are always free to speak up openly in the meeting.

You can go to your co-chairs who will consult with NQF staff or you can talk to NQF staff directly.

Any questions? Anything you want to discuss?

Okay, thank you.

DR. WINKLER: Okay. Before we get started with looking at the measures, we do want to talk about NQF's -- your role and also the portfolio of measures in this topic area.

The role of the committee, you all have varied backgrounds and we are going to ask you to wear multiple hats while you're here.

This committee is looking at a widely

disparate topic areas, eye care, ENT, don't really seem to have a whole lot in relationship to each other except historically when both groups of specialists belonged to the same professional society. So, there is a historical precedent for putting these two together.

And so, some of you may have expertise in, say eye care, but not so much the others. So, when we're talking about the measures in your area of expertise, that certainly is the hat we're expecting you to wear.

However, as a multi-stakeholder group, you also are able to wear several other hats. So, if eye care is not your specialty, you certainly can wear the hat of a potential patient or family member.

Most of you are involved in the clinical care of some type of patient. So, as a professional, an interest in moving towards a higher quality, more efficient health care system is something that we should all have a stake in. And so, you wear that stakeholder hat as well.

And so, I'm asking all of you to stay engaged in all of the measures and offer what you're able. So, if you don't necessarily have that clinical expertise, put yourself in the position of a patient, a generalist or a member of the community in terms of asking how does this measure going to help drive improvements in the quality of care provided in the health care system.

So, again, remember that you're representing a wide variety of interests and stakeholders. And so, depending on the topic at hand may very well be you're wearing different hats today.

So, I really appreciate your thinking in those terms.

As I mentioned, you will be assigned by a draw. We're going to do it, I think, after lunch today, a two or a three year term of office for a Standing Committee.

In the off time when we're in actively engaged in a project, we may come back to you

with questions that come up around some of the measures.

We do annual reviews of measures. Sometimes we need to do ad hoc reviews of measures. Sometimes there are issues that arise in a topic area and we may be coming back to you and asking you for your input.

So, it will be less busy than now, but we certainly anticipate needing to touch base with you over that time frame until the next big review of measures with other questions.

Your role is with all of us to achieve the goals of this project which essentially are to look at the measures. Most of the measures that we are going to be looking at are currently endorsed by NQF and so we want to look at continuing endorsement and whether they still meet NQF's criteria.

We need to evaluate each of the measures against all of the criteria. After your recommendations, we will publish them and seek public comment to get feedback on your

recommendations.

We really enjoy getting the feedback to find out what's happening in the field, what folks -- the audiences out there think of the measures, whether it's practicing clinicians or other groups.

And also, the CSAC, the Consensus Standards Approval Committee, which is the subcommittee of the Board of Directors is sort of the end of the road when it comes to granting endorsement along with the Executive Committee of the Board. And so, they may as any specific questions or directives. And so, it's your job to respond to those as needed.

So, we will be guiding you through all of this process. But, again, we do appreciate your being here.

Does anybody have any questions on that before we move on?

Okay. Okay. In terms of the measure evaluation which is our primary activity for today and tomorrow's meeting is that, as I

mentioned, all members are expected to be involved in the review of all measures. We have gone through in the workgroups the discussants. We've sort of had a run through.

We will evaluate each measure against the criterion. This will be a formal process of going through each on, discussing it and the committee voting so that we do have a record that can be part of the consensus process.

You will be ultimately making recommendations on these measures for continued, or for the one new measure, endorsement in an ongoing way.

And then we're going to discuss in a few minutes the overall portfolio of measures. You're looking at the vast majority of them for review today.

However, we would really like to see your input and thoughts of where there are areas that measurement, we don't currently have measurement gaps, things that would be useful to drive further improvements. And we're looking

for your input. So, those'll be some of the conversations we have tomorrow as we finish up evaluating the measures.

Okay, so, I just want to briefly talk about the portfolio. And, as I said, we put eye care and ENT together. From an NQF perspective, operationally, because they're relatively small groups themselves but they do have a historical relationship. So, we're sort of relying on that.

But, as you have probably would do, looking at it from the perspective of eye care and ENT.

So, the next slide? Whoever's over there? Okay.

So, the eye care portfolio, and I just want to look at it in its toto and you will see there two measures for macular degeneration that you're going to talk about, three measures for cataracts, two measures for diabetic retinopathy and then I've added one measure that is not primarily the responsibility of this committee but is very much related, and that is Measure 55,

which is a measure for patients with diabetes that is overseen on our endocrine committee but it is the referral for an eye exam.

And so, I want you to be aware that that measure is also part of NQF's overall portfolio.

Then we do have two measures for glaucoma and we do have a new measure around vision screening that we'll discuss.

I'd like to point out something. We've endorsed many of these measures for a long time. I think I've mentioned to you all that I've been with NQF for more than 14 years and so, I do have the benefit of history and knowing.

I'd like to particularly point out some of the measures in the eye care portfolio as some of our earliest outcome measures that came through. In the early days, process measures was the main focus and there was both a technical kind of wariness of outcome measures, but we really have had some outcome measures in eye care from some of the earliest times and I do

congratulate the developers and the society for pushing in that direction.

But I'd like to specifically point out the three measures for cataracts as a particularly wonderful group of measures because what you have is patients have undergone cataract surgery, 1536 is how the patient thinks about their vision improvement after 90 days. So, it's an outcome based on what the patient perception.

Second is, you know, clinical assessment of, you know, whether their vision is 20/40 or better afterwards, so there's an objective assessment.

And then there's a measure of complications.

So, we have three very too the point outcome measures that pretty much describe what happens for patients undergoing cataract surgery.

So, this is a particularly nice grouping for this subject area. And we do, NQF does put a premium on outcome measures and we certainly look for outcome measures whenever it's

appropriate.

We certainly are looking forward to seeing more patient reported outcomes such as we see here. And so, your input on the types of patient reported outcomes that might be particularly pertinent in these topic areas would be very useful.

So, this is our eye care portfolio as it is and we will be asking you to help us talk about the measures in the portfolio as well as opportunities to improve the portfolio.

The next slide is the other side of the coin which is the ear, nose and throat conditions. And, again, we're going to be talking about the group of measures around ear infections and one measure for pharyngitis.

I did want to point out that there is another measure that was discussed as related to this measure for treatment of children with a URI, so it is related, although it is managed by another Standing Committee in our pulmonary topic area.

And then we do have several measures for speech and hearing, so the functional sense is there.

So, that's the portfolio. And so, again, your input in terms of other topic areas, where there are gaps, how we might move to outcome measures, where there are opportunities for patient reported outcome measures would be quite valuable in trying to, you know, understand where this portfolio could be improved going forward.

So, any questions about the portfolio? Because one of your roles is to oversee it.

Okay, go on.

So, I just want to let you know that there were several measures as we came to bring these topic areas to you.

Three measures, the measure developers decided not to continue to pursue endorsement, so we really want to just be complete and share with you the measures that will be removed from the portfolio because they've been retired by the

developer.

Next one? Okay.

MS. LUONG: Hi everyone.

So, I'm just going to briefly go over the projects, activities and time line with you all again. I know I mentioned it during the orientation call, but just to remind everyone.

After the in-person meeting, there is a post-meeting webinar to go over anything that we did not cover at the in-person meeting today and tomorrow and that is on June 22nd.

Staff will then begin to draft the report. We'll begin shortly after this meeting as well as with the post-comment, post-meeting call on the 18 measures in this project. And the draft report will go out for commenting on July 10th to August 10th.

During this time, member of the public as well as NQF membership can comment on any of the measures in this project.

The comments will be gathered by NQF staff and we will present you with the comments

at the August 21st meeting which is the post-comment call.

And during this meeting, we'll present you with the comments, the themes for the comments as well as your proposed responses to the comments.

NQF staff will then gather your suggestions and comments during this call and we will put it back out in a draft report with red lines and it will go out to membership for voting and that is in September, from September 9th to September 23rd.

So, that happens. And after NQF membership votes, we will bring all of this to the Consensus Standards Approvals Committee, also known as CSAC. And this is the next body to oversee the measures and they have a few options to vote.

They can vote to uphold your recommendations of the measures. They can vote not to or they vote if they have concerns with the measures for it to be brought back to the ENT

Standing Committee for further discussion.

And CSAC is expected to review the measures in this project on April 13th. And once CSAC reviews it and votes, it goes to the Executive Committee of the Board of Directors for the final measure ratification. And that's on November 13th.

Once the Board ratifies the measures, it goes out for appeals for a month and that's from November 19th to December 18th.

And we expect to have the final report published on the NQF website early in next year.

So, I will now give this off to the two co-chairs to discuss some ground rules for today's meeting.

CO-CHAIR MERENSTEIN: Thank you.

So, this is our first meeting doing this, being at any of the meetings, our being chairs, so hopefully, you can help us and move it along.

We'll try to stay on time. And the few rules are on the slide but be prepared, and

having reviewed the measure before hand and everyone's already participated in the phone call so I think we're okay with that.

Try to base evaluation recommendation on the measure evaluation criteria and guidance. So, obviously, we're going to bring in other things that just from your baseline knowledge, but try to use that and just point out when you're bringing different things in.

Remain engaged, Reva already talked about that. There's lots of ones that, you know, are not really pertinent to what you do daily, but still try to remain engaged.

And then if you guys need a special break that's not on here, just let us know. But otherwise, we'll just stick with the breaks that we have set up.

CO-CHAIR YAREMCHUK: And another point is to keep it concise and focused. And often times, many of us who've done research in areas and we may tend to get off on a tangent on something or another. But, remember, we're

looking at the measures and the applicability of the things that we're interested in.

And if things have been said previously, it's okay to say agreed and then just move along.

It's important to think about your different constituents and your constituents are your patients, your colleagues and people that will be using these measures.

And then, what Reva had talked about earlier was the idea of improving quality of care. You know, underutilization, overutilization and being appropriate for our patients.

And I think the indicated agreement without repeating what has been said, our first measure going through this because we're a group that hasn't worked together before may be a little bit more structured in terms of going through it. And by the end of the day on Thursday, I'm sure it'll be a little bit faster.

So, it's a pleasure to have everybody

here and I think Dan and I both recognize, and Reva and everybody from NQF, the amount of time it takes and energy and away from business. And often times we say we have a day job going on as well and so, trying to stay here and having your day job at the same time can be challenge.

MS. LUONG: So, thank you both, Kathy and Dan.

Now, just to go over the process of the measure discussions, I'd like to go over how it will flow.

You have in front of you in the packet a measure discussion script. This will help you with facilitating -- with understanding how the flow of the measure discussion will go. And I'll also speak a little bit to it now.

As we mentioned at the orientation call, NQF continuously strive to improve our committee meetings based on input from a variety of different stakeholder groups.

And we are very fortunate to enough to have our measure developers present at the

meeting, especially today with a full house in the back. And we will be asking them to briefly introduce their measures as their measure goes up for discussion.

And in the case of our measures, since they are grouped in many topic areas, they can give a brief introduction of the measure groups.

So, committee members are -- after the measure developers discuss their measures briefly for two to three minutes, the committee members who are the discussants on the measures will then begin to discuss the measures in relation to the NQF measure criteria.

We have provided a designated place for the developers on the two sides to my left right now. And they will be able to speak during their measure discussion and it allows them to answer any questions that you may have for them at a more close setting.

Both committee members and developers may put their name cards up, the name cards right in front of you, when they wish to respond to a

question raised or to correct any statements that may come up about their measures.

For those on the webinar, Vaishali and Judith, you can -- there's an option for you to raise your hands if you'd like to speak or you can just speak to the phone.

During the measure evaluation, committee members often offer suggestions on how to improve the measures and these suggestions can be considered by the developer for future improvements.

However, the committee is expected to evaluate on the measures in front of them and make recommendations per the submitted specification and testing.

As Reva mentioned earlier, the committee members, you, act as a proxy for the NQF membership. And, as such, you, as a multi-stakeholder group, brings together a variety of different perspectives and values and priorities to the discussion and that's one of the things that we are very fortunate about.

It is very important that we respect each other's differences of opinions and stay friendly and collegial with our interactions amongst each other and as well as with the developers.

So, moving forward, to recap the voting criteria, there are four to consider today for each of the measures discussed. During the discussion of the different criteria, we'd like it if you can stay within that discussion.

So, for example, if we are discussing performance gap within importance to measure and report, the first criteria, please stay within that discussion and not diverge to maybe usability and use. We will have time to talk about that later.

So, the criteria are in specific order and there is a hierarchy, there is a logic to looking at them in this specific order.

The first one is important to measure and report. And there are two subcriterions under it which is evidence and performance gap.

Following that is scientific acceptability of the measure's properties and looking at -- this is looking at the specifications as well as the reliability and validity of the measure which are the two subcriterion.

And it's important to note that the first two criteria and the subcriterion under it are must pass criteria.

Moving forward, criteria three is feasibility. And the ideal goal here is to have a measure that causes as little burden on the end user as possible.

As for criteria four which is usability and use, the criteria is really on how the measure can be used for decision making regarding accountability and improvement later on.

And now, I am going to hand this over to Kaitlynn who will be going over our voting process and how you can use the clickers right in front of you.

MS. ROBINSON-ECTOR: Hi everyone.

So, just to go over this quickly. In order to vote, you point your clicker directly at this laptop and each number on your clicker correlates to an option on the voting slide. And if you wish to change your vote during the voting process, simply click the number that correlates to the new answer that you would wish to put in.

If we stop the voting and you would like to change your vote, then just me know and I'll open voting again.

Thank you.

DR. WILSON: The other thing I would mention is, your clicker is a unique number, so it's not like Chicago where you can vote early and vote often. So, don't worry if you change your vote and sometimes we may miss a vote and we don't get to the number we're expecting so we'll say vote again. So, don't worry, you're not putting in a duplicate vote. So, just know that your clickers are unique.

And, yes, sir?

MEMBER SCHACHAT: Could you explain if something passed if there's a majority or do you need 80 percent or what the cutoff is?

DR. WILSON: We're going to get there. Is that in your slides? Do you want to cover that?

MS. LUONG: Yes, I can briefly go over that. It's on the voting slide it has the consensus on it.

So, overall, we have 16 Standing Committee members, one cannot attend today and two will be attending by phone.

And so, in the room we have 13 people. Quorum for a Standing Committee is at 66 percent, so we have more than that right now in the room, especially with the two on the phone.

And in terms of consensus not reached, it is anywhere from 40 to 60 percent.

So, for the two must pass criteria, if we somehow end up in the gray zone, consensus not reached for this, we will continue to move on and we'll make a note of it in our draft report.

For measures, for the two, especially the criteria one and criteria two, if there are less than 40 percent of vote in favor of the measure, then it does not pass. And if there's more than 60 percent then it's a pass.

Do you have any questions?

MEMBER LYNCH: This is Judith. I'm still having problems with figuring out how to vote. I do not see any clicker on the up side of the left on anything that you have sent me.

MS. LUONG: Sure, Judith, yes. That's another thing that we're working on. But, in the meantime, for Judith and Vaishali, can you email me your voting results? Or, if you feel comfortable, you can just -- yes, so just please email me your voting results and we'll speak to the voting options.

MEMBER LYNCH: Thank you.

MEMBER PATEL: Yes, and that's fine, right.

MEMBER SCHACHAT: So, to reach consensus it's got to be 60 percent of 15 because

we have 13 here and two online?

MS. LUONG: Oh, and another thing, so, I know I mentioned this earlier, but if you'd like to speak since we're all here and we might want to say something at the same time, please use your name card.

And also, for the microphones, when you're done speaking, please just turn it off because there can only be three microphones on at the same time.

Thank you.

DR. WINKLER: Any other questions about the voting process?

As we do the first measure, we'll all have a chance to kind of go through it.

Is that the last slide, guys?

MS. LUONG: Yes, that is.

DR. WINKLER: Okay. All right, so I think we can turn it over to Dan and Kathy and we can get started with the first measure, 0653. So, if the measure developer representatives want to join us at the table?

MS. GORHAM: Before we get started, Vaishali and Judith, this is Shaconna. Did you receive my email?

MEMBER LYNCH: Yes.

MS. GORHAM: Okay. Can you access the path?

MEMBER LYNCH: I can.

MS. GORHAM: Okay, wonderful. Vaishali, are you okay as well?

MEMBER PATEL: I'm just checking. Oh, yes, yes.

MS. LUONG: Oh, so you can see the chat now? Everything is good?

MEMBER PATEL: No, I can't see the chat. All I have are the measures.

MS. GORHAM: They have the measures.

MS. LUONG: Okay, so if you can just email me when we vote on each criteria, that would be great.

MEMBER PATEL: Okay.

MS. LUONG: Thank you.

MEMBER PATEL: Thanks.

CO-CHAIR MERENSTEIN: So, we have two measures with otitis, so we're going to -- I think the developers are going to go over both at the same time. Is that right?

DR. ROSENFELD: Good morning. I'm Rich Rosenfeld. I'm a pediatric otolaryngologist at SUNY downstate in Brooklyn.

I've been involved with both creating the guidelines as well as the performance measures that we'll be discussing today. And I appreciate the opportunity to be here.

Thank you.

I have no conflicts of interest even though I'm -- other than I like the measures and the guidelines, but that's the conflict.

So, the first two are on acute otitis externa, affectionately known as swimmer's ear, not so affectionately known by the people who get it because it ruins your vacation and life very quickly.

It's defined as a diffuse inflammation or infection, really a cellulitis of the skin of

the ear canal. It can involve the ear drum as well as the skin of the external ear.

You know you have it because it hurts, you get itching, you get fullness, but most importantly, it is extraordinarily, sometimes excruciatingly, painful to the point that it's incapacitating and requires narcotics.

The reason we're interested in this from a guideline perspective and a measure perspective is it's very, very common, there's about 2.4 million annual visits in the U.S. every year. It affects roughly 1 in every 120 in the U.S. population and the lifetime chance of getting it is about 10 percent.

It's remarkably common. It's also remarkably mismanaged which is why we feel there's an opportunity for quality improvement.

The two measures that we have up, one relates to -- both relate to process and they relate to doing things in the first case that you shouldn't do. It's an overuse measure of antibiotic prescribing that the clinician should

not prescribe antibiotics for treating this condition, oral antibiotics that is.

There are two randomized trials that have looked at this and found no benefit when you add it above and beyond the topical drugs. But most importantly, the main organism that causes swimmer's ear is called pseudomonas aeruginosa and the vast majority or oral antibiotics are completely ineffective against this.

CO-CHAIR YAREMCHUK: Rich, we're talking 0653, so this is the topical prep -- I know, but this is the one that we are going to try to --

DR. ROSENFELD: I was told to introduce both --

CO-CHAIR YAREMCHUK: Okay.

DR. ROSENFELD: -- measures at the same time --

CO-CHAIR YAREMCHUK: All right.

DR. ROSENFELD: -- unless I'm doing something wrong.

CO-CHAIR YAREMCHUK: All right, and I

don't know which order that you wanted to do it.

DR. ROSENFELD: Okay.

CO-CHAIR YAREMCHUK: Okay.

DR. ROSENFELD: Well, does it matter which one goes first in the introduction? We'll do it on your schedule.

DR. WINKLER: We'll do it on the order, but your introduction can encompass both --

DR. ROSENFELD: Okay.

DR. WINKLER: -- to quickly get them into the picture.

DR. ROSENFELD: Sure, no, it'll be brief.

So, using an oral antibiotic to treat this not only doesn't work, it's actually a nocebo effect because it has lots of adverse events.

Roughly about 31 percent of providers seem to be doing this despite a prior guideline and the known lack of efficacy.

The second measure, or actually the first measure we'll speak about relates to

topical therapy for acute otitis externa which is supported by about 31 randomized control trials as well as several systematic reviews including a Cochrane review, all of which show very good efficacy.

About two-thirds to 90 percent of people, 95 percent of people are going to resolve in 7 to 10 days using these products. So, the measure encourages the use and promotes appropriate use by asking people if they are indeed using the topical preparation.

The topical preparations are not just antimicrobials, they can be antiinfectives in some cases, most commonly, acetic acid.

Even when it is an antibiotic, the risk of adverse events is extraordinarily low mainly because it's like dropping an atomic bomb on the bacteria it's so concentrated that there's really no resistance, similar to using ophthalmic topical preparations. It's that old saying that, I guess, dead bugs don't multiply. So, it's very effective.

The profile of safety is extraordinarily good and it tends to be underused. We don't have exact data on this, but there is clearly a gap here that people don't use them or use systemic instead.

So, that's the introduction. And Kathy and Michael want to talk about the evidence?

CO-CHAIR YAREMCHUK: So, this was an existing measure that was accepted and endorsed in 2011 and so it's coming back again for updating.

There was a clinical practice guideline that was issued by the American Academy of Otolaryngology. So, there was a review 2014, so the background and the history of this is early, fresh and not outdated at this point in time.

And there was also a Cochrane Collaborative review on this as well.

So, there's been two systematic reviews that endorse, and I'm going to talk about

the topical therapy measure first.

And so, once again, it's a process measure. It is something that can be gotten from administrative claims, administrative claims being that you have an office visit with an ICD-9 or an ICD-10 code for otitis externa.

And then there is a pharmacy claim for topical therapy.

So, it is very a common problem. It tends to occur 50 percent more in the from 2 years to 14 years of age, a little bit heavily weighted in terms of pediatrics.

At the same time, almost everybody in the room will have gotten one or has had one. Once again, it's the swimmer's ear diagnosis. So, good evidence for this.

CO-CHAIR MERENSTEIN: I think we stop at the evidence unless Michael has something to add about the evidence.

No? So, then we vote. Do we have to vote on the evidence?

DR. WINKLER: We first can see if

anyone else has questions or comments.

CO-CHAIR MERENSTEIN: Oh, anyone else have questions about the evidence? It's pretty straightforward, I mean. Comments? Questions?

DR. WINKLER: So, we'll give the voting thing a try. You get to play with your clickers.

So, in terms of voting, just to be sure, the evidence, you'll rate as either high, moderate or low, depending on how you evaluated the quality, quantity and consistency based on your discussions.

So, go to the next slide for evidence. Okay. Next one? It's not an outcome. Kaitlynn? Okay, there we go.

MS. ROBINSON-ECTOR: Yes, so if you can just point your clickers at me and -- yes.

So, voting for Measure 0653 is now open. Oh, yes, and for those on the phone, option one is high, two is moderate, three is low and four is insufficient.

And we're just waiting for the votes on the phone now.

I'm sorry, so if everyone in the room could point at me one more time and press their option, we're missing one.

Great, so all the votes are in and voting is now closed.

Eighty-seven percent voted high, 13 percent voted moderate, zero voted low and zero voted insufficient.

So, for the criterion evidence, the Measure 0653 passes.

MEMBER STEIN: So, now we're going to go on to opportunity for improvement.

CO-CHAIR YAREMCHUK: I had mentioned from 2011, it is a PQRS measure, so an individual or a group can report.

And currently, from 2012, 83.9 percent of individuals that report the measure do systemic issues, they do topical antibiotics for otitis externa.

So, we're looking at that there is a gap of about 16 percent individuals in terms of reporting that. And so, 16 percent may not seem

like a huge opportunity but when you look at the volume of 2.4 million visits for this, clearly, there's an opportunity in terms of doing better.

CO-CHAIR MERENSTEIN: Michael, do you have anything to add?

MEMBER STEWART: No.

CO-CHAIR MERENSTEIN: Any questions about the improvement there?

There's a question. Rich?

DR. ROSENFELD: Since the PQRS is voluntary and it's probably primarily otolaryngologist, the actual gap is probably much higher than the 17 percent, although we don't know for sure.

CO-CHAIR MERENSTEIN: Can we open the voting?

MS. ROBINSON-ECTOR: Yes. So, voting for Measure 0653 is now open and voting will close when we reach 15 votes.

And for those on the phone, option one is high, two is moderate, three is low and four is insufficient.

Okay, all the votes are in and voting is now closed.

Seventy-three percent voted high, 27 percent voted moderate, zero voted low and zero voted insufficient.

So, for the criterion performance gap, the measure passes.

CO-CHAIR MERENSTEIN: We're going to go on to how reliable the testing is.

CO-CHAIR YAREMCHUK: Okay. In terms of reliability, once again, we're going back to what's been reported so far with PQRS. And this is based on CPT 2 Codes, ICD-9 Codes, so an electronically administered data pull which I think is reliable.

When you look at reporting of the people that have done this for PQRS, there's 11 percent of individual clinicians that have complete data and the minimum number of patients. Twenty-two percent of groups had complete data.

I think this is something that, once again, is good enough, adequate in terms of

reporting and important enough as being able to pull the data that it can move forward.

CO-CHAIR MERENSTEIN: Michael?

MEMBER STEWART: There's very little overlap in codes or any other issue here. This is a pretty straightforward diagnosis. The straightforward ICD-9 Codes, I think the data should be very reliable from electronic pulls.

CO-CHAIR MERENSTEIN: I just have a question of clarification.

So, you said it was not required to report this, so how does this get reported? I'm confused about that.

MS. TIERNEY: Yes, so, the PQRS program is a voluntary reporting program. Up until 2015 it was an incentive-based program, so incentives were available for those who reported.

In 2015, it's moving into a penalty phase so, I imagine -- and we've seen actually the numbers of reporting providers has increased over time. So, I think in 2007, it was something around maybe 17 percent. Does that sound right?

And I know in 2013, there were 51 percent of eligible professionals who were reporting in the PQRS program.

But within the program, there are a number of measures, probably close to 300 measures. And so, any individual physician or eligible professional can pick from among those measures to report.

So, these two AOE measures are just one of those, or two of those 300 that could be reported on.

So, it's really up to the physician or eligible professional to determine what patient population may typically treat and if these measures would fall into that and if they'd like to report on it.

So, it's very up to the individual eligible professional to determine what they would like to report on.

MEMBER FRIEDMAN: But just to be accurate, it's still voluntary. It's gone from --

MS. TIERNEY: It's still voluntary, yes.

MEMBER FRIEDMAN: -- incentive to penal, but it's still voluntary.

MS. TIERNEY: Yes, absolutely, yes. Thank you.

CO-CHAIR MERENSTEIN: And it sounds like it's voluntary. You get to pick which one you're good at. You perform which one you're good at, right?

MS. TIERNEY: That's probably what people would do, yes.

CO-CHAIR YAREMCHUK: Well, and the only other thing to clarify, it really -- you don't get paid whether you have zero percent or a 100 percent. You get paid for reporting.

MS. TIERNEY: Yes, absolutely.

CO-CHAIR MERENSTEIN: Andrew?

MEMBER SCHACHAT: Thank you.

I understand the ICD-9 Codes specifying the diagnosis. I don't understand how you know if a drug was prescribed from the

administrative claim status. If you could explain because I'm eager to look at the pharmacy plan data set. And then what about patients who don't have a pharmacy plan and are just getting a prescription?

MS. HANLEY: So, the data that our testing has come from is the PQRS data. So, this is data that gets reported on the claim. But the physician actually puts it's called a CPT 2 Code. There's a code that indicates the topical preparations were prescribed. So, that code gets added to the claim, gets into the process and that's what's analyzed.

MEMBER STEIN: This may be a naive question, but are some providers prescribing both topical and oral and how does that -- how do your measures deal with that?

DR. ROSENFELD: The answer is yes. I don't have an exact number but it's probably about a third who give both. And that's why we have two distinct measures, one relating strictly to the systemic antibiotic and a separate one

relating to the topical antibiotic -- topical preparation.

MEMBER STEIN: But if the goal is to get people to use the topical, why not just make one measure that gets to that?

DR. ROSENFELD: I think we could, but if -- I think the goal is twofold. It's not just to encourage topical, it's to strongly discourage systemic as well. And unless we have some way of monitoring the use of systemics, someone -- if we just add this measure, for example, the patients who are prescribed the topical, if the clinician gave both, the topical and a systemic, they would meet the measure, even though that's bad practice.

So, I think you do need the two separately to really drive down the use of the systemic and drive up the use of the topical.

MEMBER STEIN: So, systemic only is worse than nothing?

DR. ROSENFELD: Yes, it's essentially a nocebo because the most systemic oral

medications are not going to cover Pseudomonas with rare exceptions, especially in the younger population. In adults, if they're given a quinolone, perhaps it would have some benefit, but for children, there are no FDA approved commonly available antibiotics that would cover Pseudomonas. So, it would almost certainly strictly be a nocedo in that situation.

MEMBER GOLDBERG: I think he's concerned about exceptions as well. Can you elaborate on that?

DR. ROSENFELD: Can you be more specific with the question about exceptions?

MEMBER GOLDBERG: The exceptions for the use of systemic antibiotics for otitis externa?

DR. ROSENFELD: The clinical perspective, and I know our AMA colleagues can give more of the methodology, but there are exceptions that we specify both in the measure and the guideline, the main being that if the acute otitis externa spreads beyond the confines

of the ear canal to start creating the cellulitis of the skin of the external ear or even into the face or neck that that is a situation where a systemic antibiotic would be appropriate.

In very rare cases where you simply cannot get a topical medication into the ear because of debris that cannot be removed or because of edema that prevents access, a systemic drug could be necessary.

And in certain individuals with immune deficiency, diabetes or they're on chemotherapy, other conditions that make them more susceptible to rampant infections, we do acknowledge that a systemic drug would be indicated.

Do you want to add to that?

MS. TIERNEY: Yes, so to add to what Dr. Rosenfeld said, the measure is -- both measures were designed to be sensitive to various indications in which either topical therapy in this case would not be appropriate from a medical reason or a patient reason, patient refusal.

And then they also -- the other

measure just includes medical reasons in which you might want to prescribe systemic antibiotics.

So, the measures were designed to be sensitive to that. I will say they are designed in accordance with the exception methodology that's been developed by the AMA Physician Consortium for Performance Improvement.

And so, you've probably noted that the exceptions are broad in nature. So, there's three technical buckets that you could put patients in, medical reasons, patient reasons or system reasons.

And for any given measure, we would work with the expert workgroups to determine whether or not those might be appropriate for that individual measure.

So, in this case, for the topical preparations measure, we have medical reasons and patient reasons. And the reason we've done that approach is to allow for individual decision making appropriate to the needs of the individual patient and to allow for some clinical discretion

as appropriate for individual patients.

CO-CHAIR MERENSTEIN: Andrew and then Scott?

MEMBER SCHACHAT: Do you have to click a button that says I prescribed the topical to populate that CPT Code or do the electronic medical records know that if you wrote a prescription for it, it clicks that for you? Because otherwise, you'll under ascertain.

MS. HANLEY: So, this measure specified for claims reporting. So, it would actually require that the practice indicates on the claim that that was prescribed. So, it's --

MEMBER SCHACHAT: I wondering how automatic that is? Do I have to click the CPT Code as part of the visit or if I write a prescription, do electronic medical records sort of do that?

MS. HANLEY: The CPT Code would need to be reported. Individual practices may set up systems within their practices to help automate that.

MEMBER SCHACHAT: The other side of it, on the systemic measure looking to see that you don't use the systemic medicine, if you prescribe one, do you have to click a button that says you prescribed it? If you don't click that button, then you can get away with prescribing it and you won't be scored as having given it. Right?

MS. HANLEY: For the other measure, what is calculated as meeting the measure is whether or not it was not prescribed. So, again, the presence of a prescription would indicate it was prescribed and not meet the measure.

MEMBER SCHACHAT: No, there was a prescription for that one, okay, whereas you're telling me it doesn't know if there's a prescription for this one?

DR. WINKLER: Let's not confound the two measures. Let's stick to the topical right now and then we'll talk about very good questions that should be addressed.

MEMBER FRIEDMAN: Yes, I guess Andy and I think alike. So, again, you can prescribe

topicals and systemics and that's not best clinical practice and you report on topicals so you get rewarded for that but you don't dinged for giving systemics. And how do you get around it?

MS. TIERNEY: So, the measures are intended to be reported together. They are paired and that's -- we've submitted them to NQF as paired measures.

And, in fact, for the PQRS program have indicated that they should be reported together because of that exact reason. We want to make sure that, you know, both measures together, as Dr. Rosenfeld explained --

MEMBER FRIEDMAN: So, how do you know?

MS. TIERNEY: I mean that's the intended --

MEMBER FRIEDMAN: You should, I agree, you should.

MS. TIERNEY: Yes.

MEMBER FRIEDMAN: But people are gaming -- people theoretically could be gaming the

system.

MS. TIERNEY: I mean I would say there's only so much you could do as a measure developer in terms of encouraging implementation to adhere to your intended use of the measure.

And so, we've certainly indicated that in the PQRS specifications and we've indicated that in the NQF submission. So, anyone pulling the measure off the NQF website would see that it is paired. The two measures are paired. So, we would ideally implement them together.

But I suppose it's possible that that might not happen. But that's not consistent with the intent or even the way they may be endorsed if they were to get endorsed here going forward.

MEMBER STEIN: Just to follow-up on Andy's point. I think that linking it to the actual prescription of the medication would be a lot stronger than checking off if someone did or did not do it using the indirect CPT Code.

CO-CHAIR MERENSTEIN: Can I just go back to the exclusions? It seems like you're

being a little too lenient.

If it spreads outside the ear, then you have a new diagnosis called cellulitis. And just because the patient wants it, I don't think that would qualify it. I mean it's clearly indicated, as Rich explained very well, there's no indication for oral antibiotics.

So, even if the patient wanted it, I don't think that should be a reason for exclusion for doing clearly the wrong thing.

MS. TIERNEY: So, our measure development workgroup agreed. So, for the second measure, there is only a medical reason exception. There's no patient reason exception.

So, a patient, even if you have a patient who asks for it, you would technically fail the measure because there's no way to account for that within the measure for the systemic antibiotic measure.

But, for this measure, there are medical reasons and patient reasons for why a patient -- yes, please, go ahead.

DR. ROSENFELD: So, just for the topical, for example, there are certain, you know, elderly populations that are going to be unable to, because of dexterity or, you know, the environment they're in, administer a topical preparation.

Some children just simply will not let their parents put something in their ear. You literally have to press them down with velcro and have two people hold the head.

So, I think the intent here is just that if there is some patient-based reason that there's no way you're going to get that topical in that we don't want to sort of ding the person because of that.

CO-CHAIR MERENSTEIN: I guess that makes sense, though you did say it was a nocebo. So, if, you know, if they're doing oral, they're probably doing more harm than good.

MEMBER FRIEDMAN: Yes, I'm new to this process and haven't developed measures. But it seems to me best practices plus topical, no

systemic, can these measures be combined into one measure and it just seems that the best practice would be percentage of patients aged two and over with a diagnosis of AOE who prescribe topical preparations and not prescribe systemic medications.

DR. ROSENFELD: I mean, I suppose you could but it would reduce a lot the sensitivity because there are two distinct actions that are being requested.

The one action is do not prescribe the systemic unless there's some very, very good reason.

And the second is, prescribe or recommend the topical which are two separate acts.

If you put them both in one, I think you lose some sensitivity there as far as being able to -- and you also lose your ability to estimate the gaps and the prevalence of what's going on with the two separate actions.

So, I think they could be combined,

but there's a lot of distinct advantages for keeping them separate. Also, the level of evidence supporting each one is a little different and the gaps, in particular, I think are something that need to continue monitoring.

CO-CHAIR YAREMCHUK: There was the question about, you know, self-reporting the CPT Codes and, you know, whether it's oral or it's topical. In most electronic health records, when you do a prescription of some sort, it goes to a specific type of code that'll tell you exactly the antibiotic and it'll tell whether it's systemic or if it's topical.

And there was one form where we looked at we saw like a thousand different ICD-9s and CPTs and that kind of thing. And I'm going to guess that in the development of this, that will include the prescription itself.

DR. ROSENFELD: I also think that this would be picked with e-prescribe.

MS. HANLEY: So, for right now, to answer that question, as we've mentioned, the

method of reporting is reporting that CPT 2 Code.

I think looking to the future, as we think about electronic specifications to query this information automatically from electronic health records, that's when we would actually develop the list of the exact antibiotics and look for that code as it gets prescribed, et cetera to populate the information for the measure.

MEMBER BRADHAM: I'm still having some difficulties understanding the gap that you referenced about when -- if you combined it, what is the gap in the sensitivity?

DR. ROSENFELD: So, if we combine them, the only thing we could measure is if a person did not give the systemic and gave the topical, which I agree: We could measure proper care. We'd have an excellent measure of proper care.

We would have a very lousy measure of poor care because we would not be able to separate out the use of the systemic antibiotics as a separate factor from the use of the

topicals.

So, it's just logistically, the two separate actions and from a logistic standpoint with the coding, with the specifications, with even the rationale and, more appropriately, the exceptions become extraordinarily difficult to conceive because the exceptions are very different for both of the actions.

So, yes, you can measure proper care, but I think you'd be giving up a lot in the ability to measure gaps to focus on the individual things and to have really meaningful exceptions that make sense.

MS. TIERNEY: If I could also just add to what Dr. Rosenfeld said, I think when you think about the, you know, the purpose of measurement being to drive quality improvement, if you combine them into one, that certainly would -- it would be more difficult to identify, perhaps, which individual action you might need to improve upon.

And so, I think from a quality

improvement standpoint, it's better to have them separated out as separate measures.

And I will say, you know, we developed a number of measures over the past ten years of our existence and there's been a lot of -- we have a couple of measures where we've sort of combined things. And so, the feedback that we get on some of those measures is that it's very challenging from a quality improvement standpoint to know where the quality gap is.

And so, I think combining the measures has a lot of benefits, both from a clinical standpoint, but also from a quality improvement standpoint -- of not combining the measures, sorry.

MEMBER FRIEDMAN: I think we should vote no. We've had a lot of discussion.

CO-CHAIR MERENSTEIN: Any other questions?

MS. ROBINSON-ECTOR: Okay, so voting for reliability for Measure 0653 is now open and voting will close when we reach 15 votes.

And for those on the phone, option one is high, two is moderate, three is low and four is insufficient.

MEMBER STEWART: I'm sorry, I should have asked this a second ago.

It's absolutely true the concern that, in fact, people can game, you know, by saying they did when they actually did something different.

But, if we're going to use that as a criteria for how reliable -- I mean, are we supposed to use that as a measure of reliability or are we supposed to assume that people are honestly reporting and then we're just trying to figure out if the reporting will tell us what the proportion of good practice is?

Because I think if we take that into account, there's a lot of measures we're going to say, well, you know, people could manipulate and lie and add codes. I mean, that's always possible. I think we have to assume that's not happening.

Or are we supposed to take the gameability into our evaluation?

DR. WINKLER: You know, the gameability aspect of it comes up a lot but, at the same time, I think you can't ignore it but I don't think it should be your overwhelming factor in considering, you know, the utility of these measures. We can't control what ultimately may happen out there in the world.

But it doesn't mean you shouldn't bring up it when you think it's appropriate but I don't think it should overwhelm everything else.

MEMBER BRADHAM: This is Tammy. Just for clarification purposes, you mentioned that this is a paired. So, does that mean that out of the 30, they have to select both of these or can they just select one of these and then select another item on the PQRS?

DR. WINKLER: Right. The issue of how PQRS implements, it is distinct from how we may look at it here. So, the developers have submitted them as paired and we will consider

them in paired.

And what that means is our expectation is that they would be used together. That's doesn't mean we have any control over what truly happens down the road and out in the field. But that's the directive we would like to see happen.

MS. HANLEY: If I could just add one thing, especially for the measures that are reported via claims. Reporting misinformation on a claim is considered fraud and so that's also just an underlying incentive to report what you do and not report something that you haven't done.

CO-CHAIR MERENSTEIN: Continue voting please?

MS. ROBINSON-ECTOR: And voting is still open.

And for those on the phone, option one is high, two is moderate, three is low and four is insufficient.

Okay, so all the votes are in and voting is now closed.

Thirty-three percent voted high, 53 percent voted moderate, 13 percent voted low and zero percent voted insufficient.

So, for the criterion reliability, the measure passes.

CO-CHAIR MERENSTEIN: We're going to go on to validity.

CO-CHAIR YAREMCHUK: What did we just vote on? Okay, all right.

So, validity testing, this has been something that's been in use previously in terms of the PQRSs that we discussed and has been reported. So, we know that it's able to do so.

And it's been in PQRS since 2009, so it's been in effect for a while.

It is electronic. It does have the ability to evaluate the care that's being delivered and that the specifications are fairly straightforward.

DR. WINKLER: One thing I just was to point out in terms of thinking about validity is that the areas around potential threats to

validity.

And I just wanted to point out some of the data that was provided around the frequency of exclusions and how frequently they are being used, that was a topic you brought up. So, I just wanted to point out that they did provide data from PQRS and PQRS GPRO. And where the overall exception rate for PQRS was about 5.7 percent of reported cases.

CO-CHAIR YAREMCHUK: And that was for individual reporting and for group reporting, it was a little bit higher?

DR. WINKLER: Right, 11 percent.

CO-CHAIR YAREMCHUK: And there was a huge variation which is what we see in everything. So, for individual, it went from 0 percent to 51 percent for exclusions and for groups it went from 79 percent to 0 percent.

So, as in anything, there's huge variation in terms of the person that's reporting in terms of exclusion.

DR. WINKLER: The other potential thing

to think about under validity is how the results spread and can this be used to make differentiation? Are there meaningful differences of the results?

And so, some data was provided to you on sort of the range and the means and this distribution of the results so you can see how well this measure may act as a tool for identifying differences among providers.

CO-CHAIR MERENSTEIN: Michael, anything to add? Any questions about this? We addressed some of this already.

No questions? We'll vote.

MS. ROBINSON-ECTOR: Okay. So voting for validity for Measure 0653 is now open and voting will close when we reach 15 people.

And for those on the phone, option one is high, two is moderate, three is low and four is insufficient.

And it looks like we're missing one vote so if everyone in the room could revote please? It looks like we're still missing

someone. Okay, great, thank you.

All the votes are in and voting is now closed.

Forty-seven percent voted high, 53 percent voted moderate, zero voted low, zero voted insufficient.

So, for the criterion validity, the measure passes.

CO-CHAIR MERENSTEIN: So, Kathy, feasibility?

CO-CHAIR YAREMCHUK: Okay. And I think we touched on this before. It can be reported from EMR as electronic submission, CPT 2 Codes, ICD-9.

And this is a question to the measure developers. If an individual wants to do a paper review, can they still submit it that way?

MS. HANLEY: For reporting in PQRS, no, that's not an option for reporting. But they could certainly use the specifications for internal quality improvement within their practice. I mean there's nothing preventing them

from doing that.

CO-CHAIR YAREMCHUK: Okay. In terms of feasibility, once again, I think it's something that would be high in terms of feasibility.

CO-CHAIR MERENSTEIN: Michael? Michael agrees.

Any questions about that? Yes, we talked a lot about this. We can vote on it.

MS. ROBINSON-ECTOR: So, voting for feasibility for Measure 0653 is now open and voting will close when we reach 15 votes.

It looks like we have all the votes in the room. Oh, yes, and for those on the phone, option one is high, two is moderate, three is low and four is insufficient.

CO-CHAIR MERENSTEIN: We still need one vote.

MS. ROBINSON-ECTOR: So, all the votes are in and voting is now closed.

Eighty percent voted high, 20 percent voted moderate, zero voted low and zero voted insufficient.

So, for the criterion feasibility, the measure passes.

CO-CHAIR MERENSTEIN: So, for the last criteria, usability?

CO-CHAIR YAREMCHUK: It should improve treatment, quality of care and prevent overutilization, appropriate utilization of antibiotics.

CO-CHAIR MERENSTEIN: Michael? Michael agrees.

DR. WINKLER: Yes, for usability and use, particularly measures that have been endorsed for a while, we really want to think about, you know, how they've been used and what the information has been used for.

Potentially, you know, what's been the impact of this measure? Have we really seen opportunity, you know, seen improvement in the quality care provided? Have we seen improving trends? I mean is this a good tool for stimulating quality improvement? That's what we're thinking about in terms of usability and

use.

CO-CHAIR MERENSTEIN: So, Richard, do you have anything to add with that?

DR. ROSENFELD: So, I agree, it's a noble aspiration. We do not have enough data over time to I think assess that. But, we're very hopeful.

CO-CHAIR MERENSTEIN: You said before it's what, 67 percent of people use topicals?

DR. ROSENFELD: I think in the PQRS it was in the 80s. It was high.

DR. WINKLER: Right, if we look at 2009 for the performance gap, 2009 it was 72 percent, 2012, it's 83. So, it's been getting use.

DR. ROSENFELD: I take it back, it's very usable.

DR. WILSON: I mean the issue is voluntary reporting. But, the voluntary reporters reported significant improvement over time.

MEMBER STEIN: Can you guys comment on the types of providers that are using this code?

Is it ENT specialists? Primary care docs?

MS. TIERNEY: We don't know. The PQRS program provides information about the number of eligible professionals. And so for this measure, it was around 85,000 eligible professionals. Of those, 3,200 reported on the measure.

So we can guess, but the number of eligible professionals, you can sort of speculate that the number of eligible professionals includes some primary care physicians.

But more than likely, given how small the number gets that probably it's primarily specialists and otolaryngologists who use -- who actually report on the measure.

But we don't know for sure and the PQRS program doesn't provide us with that level of detail related to those who report on the measure.

MEMBER STEWART: And practically, there's not a lot of otolaryngology measures to select. There's not a large menu is PQRS. There's probably a larger menu of PQRS options

for people in primary care.

So, I mean the reality is a lot of otolaryngologists are reporting because there's not a lot measures. That would be another bit of assumption that would probably indicate it's heavily otolaryngology.

MEMBER STEIN: And the people that are prescribing the oral antibiotics, are they mostly primary care docs or are they -- like, is everyone in ENT know to do things the right way and the ones that are not doing it right, people that are not ENT specialists or even within ENT there's a lot of misuse of oral antibiotics?

DR. ROSENFELD: I think the data, again, are pretty sparse. There is a UK study that looked at primary care docs and found in the 30, 35 percent range with prescribing. But I'm not aware specifically of usable information on otolaryngologists versus emergency physicians and other folks who may see swimmer's ear with some frequency. So, the answer is we really don't know.

MS. TIERNEY: Can I just add, I mean I think it's important, although we don't really know what type of specialties are reporting in PQRS, there's still a gap of those who are voluntary reporting. There's a noted gap related to the actual use of topical preparations.

So, I think that's important even though we don't necessarily know how that is. And maybe it is more otolaryngologists but it's still important, I think, to emphasize that there's still a gap among those voluntary reporting in this program.

CO-CHAIR MERENSTEIN: Any further questions before we vote?

Okay, let's open voting.

MS. ROBINSON-ECTOR: So voting is now open for usability for Measure 0653 and voting will close when we have 15 votes.

And for those on the phone, option one is high, two is moderate, three is low and four is insufficient information.

It looks like we have all the votes in

the room.

Vaishali, I didn't get your vote yet, so if you want to email me the vote, thanks.

MEMBER PATEL: Yes.

MS. ROBINSON-ECTOR: Thank you.

All the votes are in and voting is now closed.

Okay, 60 percent voted high, 40 voted moderate, zero voted low and zero voted insufficient.

Okay, so it passes.

CO-CHAIR MERENSTEIN: So, I think we just do one more final vote about suitability. Is there a discussion before that? Anyone have any final questions about this?

So, it's just a yes or a no. I guess we can vote.

MS. ROBINSON-ECTOR: Okay, so voting for overall suitability for endorsement for Measure 0653, acute otitis externa is now open.

And for those on the phone, option one is yes and option two is no.

MS. LUONG: And this is for overall suitability for endorsement for those on the phone.

MS. ROBINSON-ECTOR: Great, so all the votes are in and voting is now closed.

One hundred percent voted yes and zero voted no, so for recommendation for overall suitability for endorsement for Measure 0653, acute otitis externa passes.

CO-CHAIR MERENSTEIN: So, unless there is any further comments, we'll take a five to ten minute break and then come back and finish a couple up before lunch.

(Whereupon, the above-entitled matter went off the record at 10:31 a.m. and resumed at 10:44 a.m.)

DR. WINKLER: We had a question that may have provided some context in terms of the endorsement process.

NQF grants endorsement for three years and we try to review these measures on that sort of periodic basis. However, we do solicit annual

updates from the measure developers on all the measures to identify any updates of measures that may occur.

If there are material changes to those measures, it could prompt an earlier ad hoc review. And also, anybody could request NQF do an ad hoc review for a measure.

There have been some controversial measures. There's been new evidence come out. There's been unintended consequences out in the field. Those are the types of things that have typically prompt ad hoc reviews. So, that can be done in the intervening between the regularly scheduled review that could occur.

Again, as a Standing Committee, you would be available to do those ad hoc reviews if they come up.

So, that's kind of the implications of the endorsement. Measures that have been endorsed that you would say do not recommend for continued endorsement fall off our list. You know, how that plays out in the field with people

using them will be highly variable.

Oftentimes measures that are imbedded in programs sort of, it takes a bit of mechanics to get them out, if you will. But, nonetheless, the message is clear.

And when you do that, we really want to document the reasons why that we made that change and removed the endorsement status.

So, I know, Josh, you were the one with the question. Did I respond to your queries?

MEMBER STEIN: Yes.

DR. WINKLER: Okay.

CO-CHAIR MERENSTEIN: I was going to go right into the next discussion with Seth and Michael.

MEMBER GOLDBERG: Dr. Rosenfeld had already introduced this measures, acute otitis externa systemic antimicrobial therapy and avoidance of inappropriate use.

This is, again, a process measure and the numerator statement, the patients who were

not prescribed systemic antimicrobial therapy and the denominator was all patients two years and older with the diagnosis of acute otitis externa.

There was extensive literature showing the lack of evidence of efficacy of systemic antibiotics both from the Academy of Clinical Practice Guideline as well as the 2010 Cochrane Collaborative Review.

And so, you know, based on this, and we have two systematic reviews and we have ICD-9 and CPT Codes and, again, pharmacy claims that would allow documentation.

Plus, we have measurement of performance with continued improvement shown in PQRS program between 2009 and 2012 with an improvement in the measure from 45 percent in 2009 to 74 percent in 2012.

But there is still a gap in care, obviously, and as we have discussed previously, we're not quite sure which physicians are being measured. It appeared primarily the otolaryngologists and so --

CO-CHAIR MERENSTEIN: Seth, sorry to interrupt. We're going to vote on the evidence before we get into gap in care.

MEMBER GOLDBERG: Oh, okay.

CO-CHAIR MERENSTEIN: Sorry about that.

MEMBER GOLDBERG: Sorry.

CO-CHAIR MERENSTEIN: Michael? Do you have anything to add?

So, any questions about the evidence behind the measure?

Then we can vote.

MS. ROBINSON-ECTOR: Okay. So, voting for evidence is now open.

And for those on the phone, option one is high, option two is moderate, option three is low and option four is insufficient evidence.

So, it looks like we're missing one vote, so if everyone in the room could please revote?

Great, all the votes are in and voting is now closed.

Ninety-three percent voted high, seven

voted moderate, zero voted low and zero voted insufficient.

So, for the criterion evidence, the measure passes.

CO-CHAIR MERENSTEIN: Thank you.

Seth, we can now talk about the opportunity for -

MEMBER GOLDBERG: Okay, again, as we discussed for the previous measure, there is a performance gap. It's still significant.

The most recent reported performance result is in 2012 with 74 percent in compliance. There is a gain indicating that this has been effective but we are uncertain outside of those reporting physicians who is actually using this measure.

And so, based on the information we have, there still is room for improvement, especially based on the frequency of the condition.

CO-CHAIR MERENSTEIN: Michael? Anything to add about that?

Any questions about that? Again, we've talked a little bit about this already, but any questions?

So, let's go to voting.

MS. ROBINSON-ECTOR: Voting for performance gap is now open and we will close votes when we have 15 votes.

And for those on the phone, option one is high, two is moderate, three is low and four is insufficient.

So, all the votes are in and voting is now closed.

Eighty-seven percent voted high, 13 percent voted moderate, zero voted low and zero voted insufficient.

So, for Measure 0654, the measure passes on criterion performance gap.

CO-CHAIR MERENSTEIN: Thank you.

Now, we're going to talk about the reliability of the test.

MEMBER GOLDBERG: Okay. Again, we talked about reliability of the previous measure.

And, again, this is based on claims of the registry. The numerator is CPT 2 Code, the denominator is ICD-9 and 10 and CPT Codes.

We do have exclusions, exemptions. They are specified as an insufficiency, giving examples of insufficiency in diabetes.

The method of testing and testing results, I believe, demonstrate that this is reliable.

CO-CHAIR MERENSTEIN: Michael, anything to add?

Tammy, do you have a question?

MEMBER BRADHAM: I have two questions.

The first -- well, I think you already answered one of them, I'm sorry.

I have a second question. In the exclusions, does it include people with hearing aids?

MS. TIERNEY: So, this is Sam. So, the exclusions or the exceptions, as we call them, are broad. So, you would just put down a medical reason in something like a hearing aid if that

was an appropriate reason, you could put that down and that would be a reason that you might prescribe systemic antimicrobials.

DR. ROSENFELD: From a medical standpoint, I'm not sure that would be a valid reason unless they were so dependent on the hearing aid they couldn't function and there was a little concern about the liquid getting into the aid.

But, just the presence of an object in the ear canal, I think usually with acute otitis externa, it's so painful you couldn't put your hearing aid in and they probably would be fine with topical.

So, it's a potential exception, but probably not a particularly common one I would expect.

MEMBER LYNCH: I would agree with that. This is Judith.

CO-CHAIR MERENSTEIN: Andrew?

MEMBER SCHACHAT: So, this is my leftover question from the last one.

Is there a CPT Code that says I didn't prescribe systemic antibiotics?

MS. HANLEY: Yes, there is.

MEMBER SCHACHAT: So, it's extra work? You have to click and score another thing on the medical record, so you could easily under ascertain not prescribing the antibiotics?

CO-CHAIR MERENSTEIN: Yes, and I think Rich agrees with that, right? You think that probably it's probably higher than we're actually recording, I think that's what you said previously.

MEMBER LYNCH: So, you're saying that I would have to click that I did not order a systemic antibiotic in order to meet the criteria?

MS. HANLEY: You would enter a code on your claim indicating you did not prescribe an antibiotic for this patient.

MEMBER LYNCH: Even though I ordered a topical?

MS. HANLEY: Yes, because these are

reported separately.

MEMBER LYNCH: Okay.

MEMBER GOLDBERG: Couldn't this be picked up passively by the absence of a prescription or pharmacy report or e-prescription?

MS. HANLEY: So, that question, again, gets to the implementation of the measure and the implementation that we've used to test the measure has been the PQRS program.

And the way Medicare is designed, that program is you actually have to actively report a code.

But, again, a health system practice, they could design their own systems that would more in an automated way collect the information for the measure.

CO-CHAIR MERENSTEIN: Any further questions or comments?

So, I guess we'll vote on the reliability.

MS. ROBINSON-ECTOR: So, voting for

reliability for Measure 0654 is now open and will be closed when we reach 15 votes.

And for those on the phone, option one is high, option two is moderate, option three is low and four is insufficient.

Okay, all the votes are in and voting is now closed.

Thirty-three percent voted high, 47 percent voted moderate, 20 percent voted low and zero voted insufficient.

So, for the criterion reliability for Measure 0654, the measure passes.

CO-CHAIR MERENSTEIN: Seth, validity, please?

MEMBER GOLDBERG: Okay. In terms of validity, it does appear the specifications align with the evidence.

The measure was tested for face validity. And in terms of meaningful difference, the range in performance was quite high and consistent and it does identify meaningful differences about quality.

CO-CHAIR MERENSTEIN: Michael, do you have anything to add?

MEMBER STEWART: Nothing to add.

CO-CHAIR MERENSTEIN: Thanks.

Just as a general, when we're speaking, try to speak closer because they can't hear us back there, guys.

Any questions about that? Again, we talked a lot about this.

Let's vote.

MS. ROBINSON-ECTOR: Thanks. So, voting for validity for Measure 0654 is now open.

And for those on the phone, option one is high, option two is moderate, option three is low and option four is insufficient.

It looks like we have all the votes in the room. All the votes are in and voting is now closed.

Forty percent voted high, 60 percent voted moderate, zero voted low and zero voted insufficient.

So, for the criterion validity for

Measure 0654, the measure passes.

CO-CHAIR MERENSTEIN: Feasibility please?

MEMBER GOLDBERG: Okay, we talked about this previously. The data elements are defined in elected fields in electronic claims and the measure has been in use by PQRS since 2009.

MEMBER STEWART: Agree this is feasible.

CO-CHAIR MERENSTEIN: Any questions about that?

MS. ROBINSON-ECTOR: So, voting is now open for the criterion feasibility for Measure 0654.

And for those on the phone, option one is high, option two is moderate, three is low and four is insufficient.

So, it looks like we're missing one vote in the room. If everyone could please revote. Thank you.

Great. Okay, all the votes are in and voting is now closed.

Sixty-seven voted high, 20 percent voted moderate, 13 percent voted low and zero voted insufficient.

So for the criterion feasibility, Measure 0654 passes.

CO-CHAIR MERENSTEIN: Usability please?

MEMBER GOLDBERG: The measure is currently in use in the PQRS program which is publically reported. And there is published information that there has been improvement over time. There doesn't appear to be any unintended consequences with the measure.

CO-CHAIR MERENSTEIN: Michael, anything to add?

MEMBER STEWART: Nothing to add.

CO-CHAIR MERENSTEIN: Kathy?

CO-CHAIR YAREMCHUK: I mean the only correction, it's not publically reported. You can submit the information and you can comply with PQRS standards, but it isn't publically reported.

MS. HANLEY: If I could just add, I

would say yes. CMS is transitioning to starting to publically report some of the results on some of the measures.

CO-CHAIR MERENSTEIN: Any questions about that? I see people voting, so I guess not.

MS. ROBINSON-ECTOR: So, voting is open for usability for Measure 0654.

And for those on the phone, option one is high, two is moderate, three is low and four is insufficient information.

Okay, all the votes are in and voting is now closed.

Seventy-three percent voted high, 27 percent voted moderate, zero voted low and zero voted insufficient information.

So, for the criterion usability and use, Measure 0654 passes.

CO-CHAIR MERENSTEIN: We need a vote for overall suitability.

MS. ROBINSON-ECTOR: Okay, so voting is now open for overall suitability for endorsement for Measure 0654, acute otitis externa systemic

antimicrobial therapy.

For those on the phone, option one is high and option two is no.

CO-CHAIR MERENSTEIN: Yes and no.

MS. ROBINSON-ECTOR: Oh, sorry. Yes, option two is no and option one is yes.

MS. LUONG: Vaishali, if you can vote via email please? Thanks

MEMBER PATEL: Yes, I just did.

MS. LUONG: Okay, great. Thanks.

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

Ninety-three percent voted yes and seven percent voted no. So, for recommendation for overall suitability for endorsement, Measure 0654 passes.

CO-CHAIR MERENSTEIN: Okay, so we're going to go on to the next one on otitis media with effusion. So, if you want to pull it up and with the same developers.

Other people coming up? Thank you for your time.

DR. ROSENFELD: Okay, so I'll introduce all three of these at once, is that the plan?

CO-CHAIR MERENSTEIN: Great plan.

DR. ROSENFELD: Okay, good.

So, now we shift -- we'll stay on the ear but we switch to kids and talk about otitis media with effusion called also middle ear fluid in lay terms.

And this is defined as the presence of some fluid in the middle ear space without any signs or symptoms of acute ear infection.

And this is essentially an occupational hazard of early childhood. If you occupy early childhood, you get middle ear fluid at some point. Between 50 and 90 percent of children will get this by the time they enter school.

The period prevalence over a year, so if you look over a year, 15 to 30 percent of young children get it. And there are studies that have tracked individual children and found that young children on average will get four

distinct episodes per year.

And if you follow them over several years, they spend about 25 percent of their days, have some degree of middle ear fluid. So, this is fairly ubiquitous in your children.

The symptoms can be subtle and, at times, may be almost completely absent. The most common one would be hearing loss that can impact speech and language development and school functioning and school performance.

Other symptoms could be as subtle as a lack of feeding properly or sleep disturbance, irritability, behavioral changes, some have ear pain but it definitely does affect them in many different ways.

The vestibular problems are also seen with this where kids are quite clumsy and have gross motor issues at a young age.

We don't know exactly how many episodes there are but it's been estimated on some pretty old data that there's about 2.2 million episodes a year in the U.S. Those are

strictly new episodes. But a lot of these can be very persistent as well.

The three measures that are proposed are all very similar. They're all overuse measures and they all deal with various types of medical therapy, all of which recommend not to do it.

So, there are two, one for antibiotics, systemic antibiotics that says the physician should not prescribe it.

There's another on systemic steroids that also says the physician should not prescribe it.

And then there's one on antihistamines which says that, but also extends it to say that the physician should also not recommend anticipating the over-the-counter use of such products that extends beyond prescribing.

So they're all three process measures. We could potentially consider the one that says do not recommend is also falling into the communication category of quality improvements.

I'll just very briefly go over some of the data and gaps which we'll talk about more thoroughly with each measure.

We don't have the robust data that we had or somewhat robust data we had for the otitis externa measures because these have not been used and practiced to a significant extent. They're not PQRS. So, we have some limited data on the issues.

The antibiotics first, and the data that we do have, I will say are unpublished but fairly robust. They come to us from an individual, Jennifer Shin, who's an otolaryngologist at Harvard and has been working very fortuitously on the past several months looking at some big national databases to monitor usage of medications for otitis media with effusion.

This was motivated by the new clinical practice guideline that we're about to publish in the next few months that updates the 204 guideline on this.

So, those databases covered millions of pediatric visits in the U.S. and had about a 114,000 otitis media with effusion visits in them. So, it's a fairly significant data set.

For antibiotics, they found that about 32 percent of encounters resulted in antibiotic prescribing even though it's not recommended to be used for that purpose.

And, obviously, there are rather significant side effects and adverse effects of antibiotics ranging from individual issues, allergic reactions to the societal implications, the resistance.

For antihistamines, her work showed a 9.5 percent use of these products during otitis media with effusion visits. And, interestingly enough, and that includes -- this includes both over-the-counter as well as prescriptions, and about 60 percent of that 9.5 percent was related to the old first generation sedating antihistamines which actually do have some rather substantive side effects in children especially

since dosing is not always optimal in young kids.

The other 40 percent were more of the non-sedating antihistamines that typically require prescriptions.

And the last category of systemic steroid use is fairly low. That was 3.2 percent of otitis media with effusion visits in her analysis.

And those data, and actually, all of these data, unfortunately, do not break down individual prescribers. They're just an average of all the visits.

So, it's quite possible, and I would say very likely, that there are pockets of individual providers or group practices where there's very high prescribing.

So, for steroids, perhaps, 10, 20 percent of visits are getting this. Whereas, in many other practices, it's very low or close to zero.

So, there are probably outliers here but we just cannot identify them from this data

set. But anecdotally, it does seem that there probably are instances of abuse going on.

We'll talk about the evidence in a few moments. But, there are Cochrane Reviews on all of these which do not show substantive benefits. For the antihistamines, there's absolutely no benefits, that's old news, zero benefit, again, a nocebo strictly.

For the systemic steroids, there is a slight benefit at one month which disappears afterwards in the Cochrane Reviews, so it's not recommended.

For the antimicrobials, there is a short term benefit, actually, one that goes out a few months of about a 15 percent bump in resolution rates. However, there's absolutely no impact on hearing levels or rates of ear tube tympanostomy tube insertion which are the more relevant patient-based measures.

So, the Cochrane collaboration recommended not to use these drugs because of the lack of the really meaningful clinical impact.

So, that's the overview.

Yes?

MEMBER FRIEDMAN: I'm sorry, can you enlighten us as to the appropriate treatment for this condition?

DR. ROSENFELD: Well, thank goodness we have three measures, all of which say do nothing. But that's a good question.

Much of managing middle ear fluid in kids is sort amusing the patient while nature cures disease in the words of Voltaire. And antibiotics, steroids and antihistamines are not very good amusements because they all have side effects.

So, our new guideline, for example, is going to stress a lot supporting materials to give parents so they understand the natural history is very favorable, that most cases get better on their own.

And we also have a tympanostomy tube guideline that very clearly lays out the situations where if it's persistent for three

months or longer in both ears, particularly if it's impacting hearing or having other effects on children such as their school performance or ability to acquire speech and language, those are situations for tubes.

But it really boils down to either you're going to put the tubes in at some point or you're just leave them alone and just let them outgrow it on their own which is the more typical outcome with these kids.

MEMBER STEIN: Do you have any plan on creating a measure to capture use of tubes when appropriate?

DR. ROSENFELD: Yes. What we're going to work on probably are similar to a measure are the appropriate use criteria. We're going to derive from our tympanostomy tube guideline.

But, yes, we have a very recent tympanostomy tube guideline 2013 that has some very explicit recommendations regarding tubes and the senses. There are some real QI opportunities there.

So, we will look at measures. We'll look at appropriate use criteria.

CO-CHAIR YAREMCHUK: And this is a question. You're currently updating these guidelines?

DR. ROSENFELD: Yes. So, the otitis media with effusion clinical practice guideline was published in 2004 and that was a joint project of the American Academy of Pediatrics, the Otolaryngologist and the Family Physicians. That's where the measures stem from because we had very similar recommendations in that. At that point, they were just recommendations.

The new guideline that's in a fairly advanced for of a draft now and will hopefully be published by very early 2016 carries this concept further. It breaks it down into individual recommendations and they're all strong recommendations now against prescribing because there's been more evidence, particularly regarding antibiotics and steroids than on the first guideline.

CO-CHAIR YAREMCHUK: So, I guess the question that I have is it worthwhile to wait until 2016 when you have these guidelines out as opposed to doing it now based on the previous information?

DR. ROSENFELD: I don't think so. The thrust of the -- the measures do not change at all based on the current guidance. I think we'll be able to say perhaps there's a little more data, but those data have already been published.

The guideline itself is not the supporting data for the measures. It's the Cochrane Reviews and the randomized trials which are all freely accessible and have already been published.

MEMBER FRIEDMAN: Okay, so the next logical question is, is there any advantage to combining them into one measure? Is there any disadvantage of prescribing one versus all three? Do you care? If they're prescribing one, they're not doing appropriate care, what do you care whether they're prescribing one or all three of

them at the same time, which is probably even worse?

DR. ROSENFELD: So, yes. So, you could do a single measure and we could do, say, we could even do a single measure that says do not treat otitis media with effusion with any type of medical therapy because it fundamentally doesn't change the child eustachian tube.

The problem with that is these all have different adverse event profiles. And so, for example, steroids, systemic steroids, which are fairly infrequent at just over three percent.

The problem with those is in young kids who get this condition, there's been ample reports of disseminated varicella where kids die because they've taken steroids that they didn't need in proximity or even that they needed for asthma in proximity to a varicella or chicken pox episode and it gets disseminated. You get the aseptic necrosis. You get cataracts.

So, I think the adverse event profiles are very different. The prevalence is very

different. I think the gaps are very different. And even with though we could say the gaps are small, three percent for one and maybe just about ten percent for another, it's a very ubiquitous condition in a susceptible population of young pre-school kids.

And, to me, one of the advantages of having three separate measures is that, three years from now or whenever it comes time to renew, having these measures will allow us to really get the data we need to understand the real gaps and the real issues.

And if it turns out we're wrong and we overestimated, fine, the measure goes to sleep and disappears. But I think it's important to keep them separate.

CO-CHAIR MERENSTEIN: So, this is Seth and Michael again. I don't know if, Seth, you have anything to add on evidence?

Are we switching? Am I reading?

MEMBER STEWART: We'll switch. I'll do this one.

CO-CHAIR MERENSTEIN: Okay.

MEMBER STEWART: No, I think -- I guess I had -- we had a similar discussion on our call about this, the point that Kathy brought up, which was the guideline we're basing this on is basically a recommendation and not a strong recommendation.

But, in fact, there's been a lot of intervening evidence. There's going to be a new guideline coming out which is a strong recommendation. And so we feel the evidence is quite strong to support this, you know, since we're doing these one at a time, to support the one we're doing now which is systemic antimicrobials.

So, the evidence is strong.

CO-CHAIR MERENSTEIN: The evidence is the Cochrane Review, right?

MEMBER STEWART: Yes.

CO-CHAIR MERENSTEIN: I mean I guess there's other studies that they reviewed to make it.

Do you have a question, Josh?

MEMBER STEIN: This is more a global question. If, in general, we like a measure but there's one or two aspects of it that we don't like about it, do we have to vote the whole measure down for the developer to just change that one or two aspect? How does that work?

DR. WINKLER: We alluded to it on the introduction. We're asking you to evaluate what you have before. Your conversation and discussion and concerns that you raise is being heard by them. And so, they will take that under advisement for potential action down the road.

But we're asking you to make your decisions based on what's presented in front of you.

MEMBER STEIN: But for them to make that action, we'd have to vote it down, correct?

DR. WINKLER: Perhaps. I mean, I think each situation is different to what degree that your concern is that you would want to, you know, remove endorsement for the measure because your

concern is that big.

So, I think each situation might be different.

CO-CHAIR MERENSTEIN: Andrew?

MEMBER SCHACHAT: I just have a different question. I'm looking at a slide in front of me, not that one, and it says data source paper medical records. So, can you just remind us what the data source is?

MS. KAYE: So these particular measures, they are being submitted for paper medical record endorsement mostly because we need to get it endorsed based on the data source that we use to test the measures.

And so, for the purposes of this, we will be seeking paper medical record endorsement.

MEMBER SCHACHAT: But could someone just explain generally how you could ever do any of this with paper medical records? I mean are people supposed to Xerox their records and mail it to someone or what happens?

MS. KAYE: So, we included in the

materials, there is a chart abstraction tool. It's a bit of a worksheet and so that would guide the manual review of medical record and then you would answer the specific questions that then would feed into the measure.

MEMBER SCHACHAT: Was that to do the measure or is that just to validate that someone did it? Do they just check off a box to say I did it or do they have to fill out that worksheet?

MS. KAYE: From what I understand, it would be that to calculate the measure, so it wouldn't necessarily be the provider filling in this worksheet. It would be the person abstracting and calculating the measure.

MEMBER SCHACHAT: I'm not going to monopolize the discussion but I just am skeptical about anything that relies on paper records.

CO-CHAIR YAREMCHUK: And this is, again, just a question for the developer.

The previous two measures were electronic. They were prescription and these are

paper. And so, I'm asking why is there a difference between the two?

MS. KAYE: Well, for these measures, we were able to get electronic measures or electronic data for the previous measures because they are included in the PQRS program. So, we were able to access that data.

However, because these are not included in the PQRS program, we had to kind of rely on some earlier data from when we originally developed these measures that included a paper-based testing project. So, we didn't have that PQRS data available to us.

And I guess in terms of, to answer your question about, you know, updates or if there are aspects you are concerned about, there is also, we are hearing this and, you know, in terms of endorsement, there are also opportunities through annual updates and continued maintenance of the measure to where we could take those under consideration as well, depending on the types of concerns that you may

have.

MEMBER STRODE: Let me just follow-up on that. For whatever reason, the Feds have not put these measures into PQRS. An electronic health record could have the elements that would allow you to determine how this was treated.

A system could decide that they wanted that software to be created or a system could say, we'll go through the extensive effort required to either do a review off of the computer screen of the raw record or do a paper chart review.

MS. KAYE: That is correct. That just because we were kind of limited to paper-based testing data does not mean that it couldn't be implemented in an electronic format.

And so, there are certainly, I think, you know, in terms of the data elements, I don't see that necessarily being not possible, it's just more that we kind of, you know, due to resources, had to use the data we had available which is how it wound up in the paper-based

format.

DR. ROSENFELD: I just want to build on that because I want to be sure that I understand the process correctly.

The purpose is to vote on the measure itself, not necessarily the implementation process which is going to be quite variable.

Obviously, some things were done to collect some data to get to this point. And for the others it was electronic through PQRS. It was paper-based but that's just to get a sense of is the measure working? How is it done?

Going forward, there's no intent that this just be a paper-based measure. I mean it could find its way into however the system monitor uses it, it's corrective. Whether that was in EHR or some automatic process or any other, you know, value-based purchasing system that comes down the pike.

Am I correct with that understanding?

DR. WINKLER: Yes, you are. But, in terms of being able to evaluate what we know

about the measure for you all to make a recommendation on endorsement, we have to use the data that's available. And that's why the data source that they have tested, specified and tested, is based on paper medical records.

MEMBER STEWART: So, what this slide is telling us is, the data source was paper medical records, not that the guideline group is saying this must be done using paper medical records going forward?

MS. KAYE: That's correct. And I will add that these were in PQRS at the very beginning but due to some kind of more stewardship-based concerns, were removed. However, there are, you know, it was implemented briefly.

And we are currently working, or rather our colleagues at AAOH&S, are currently working with states to try to implement this in the Medicaid population which is a little bit more appropriate because this has a pediatric focus.

Those Medicare-based programs, you

don't really get a lot of those pediatric patients in.

And so, one of the hurdles to clear to get into these Medicaid-based programs, incidentally, is NQF endorsement. And so, that's one of the hurdles that we are seeking here.

But we are seeking avenues for maybe a little bit more appropriate implementation to these measures.

CO-CHAIR MERENSTEIN: So, this is a really good discussion with lots of different points but right now, I think we need to vote on the evidence, not on the -- so, we'll get to those points, too. But I think we should vote just on the evidence behind the test right now.

MS. ROBINSON-ECTOR: So, voting for evidence for Measure 0657 is now open.

And for those on the phone, option one is high, option two is moderate, three is low and four is insufficient evidence.

It looks like we're missing one in the room. Thank you.

MS. LUONG: Judith and Vaishali, I'm ready for your votes any time. Thanks.

MEMBER PATEL: I sent mine.

MS. LUONG: Vaishali, I actually just got a blank for you. Can you resend that? Thanks.

MS. ROBINSON-ECTOR: So, all the votes are in and voting is not closed.

So, seven voted high, six voted moderate, two voted low and zero voted insufficient.

So, the measure passes for the criterion of evidence.

CO-CHAIR MERENSTEIN: So, Michael, we're going to talk about opportunity for improvement.

MEMBER STEWART: So, I think we heard very nicely -- this, we're talking about systemic antimicrobials. We have probably the greatest opportunity for improvement because this is something that is frequently prescribed for this condition.

Certainly, the data that were presented indicated that. Certainly, that's been anecdotal experience is seeing patients who have had this condition referred in or have not infrequently still given antibiotics.

I think there is a high level of -- I think there's a high level of opportunity for improvement.

CO-CHAIR MERENSTEIN: Anything to add, Seth?

Any questions or comments about that?

MS. ROBINSON-ECTOR: So, voting for performance gap for Measure 0657 is now open.

And for those on the phone, option one is high, two is moderate, three is low and four is insufficient.

It looks like we have all the votes in the room. We're just waiting on one more. All the votes are in and voting is now closed.

Forty-seven percent voted high, 53 percent voted moderate, zero voted low and zero voted insufficient.

So, for the criterion performance gap, Measure 0657 passes.

CO-CHAIR MERENSTEIN: So, this is one we've already talked a little bit about but I don't know if you want to add for reliability now.

MEMBER STEWART: I think that this is -- I mean there is a clear ICD-9 Code. It is a prevalent condition. I think this should be extractable from EMRs and from other systems.

So, I would think that whether or not antibiotics were prescribed for this particular diagnosis should be able to be reliably measured going forward which is, I think, what we're voting on here today. Or how are we --

DR. WINKLER: You want to look at what we've been able to establish by the testing that's been done. I think, you know, it's fine to consider where that may take you in the future, but the reason that we require testing of the measure in some form is to give you some actual data to work on rather than speculate.

DR. WILSON: Well, I think the data that was collected was collected reliably and I think there's a high level of reliability based on the data that we have so far.

MEMBER GOLDBERG: I agree.

CO-CHAIR MERENSTEIN: Any questions? Comments? We've talked a little bit about it.

DR. WINKLER: Yes, reliability also includes anything about specifications, if there are any comments or questions about those.

MS. ROBINSON-ECTOR: So, voting is open for the criterion reliability for Measure 0657.

And for those on the phone, option one is high, two is moderate, three is low and four is insufficient.

All the voting is in and voting is now closed.

Forty percent voted high, 27 percent voted moderate, 33 percent voted low and zero voted insufficient.

So, for the criterion reliability, the measure passes.

CO-CHAIR MERENSTEIN: So, we're going to do validity, Michael.

MEMBER STEWART: So, this is whether the specifications align with the evidence and whether there were issues with exclusions or comparability.

I think based on what we've heard, I think the validity of the diagnosis, the treatment and the data appear to be valid. So I don't know if there's other questions about that.

MEMBER LYNCH: The exclusions only mention medical reasons, there's nothing specific.

MEMBER GOLDBERG: That was my question. Could the developers comment on that further?

MS. KAYE: So, this is a similar -- this reflects our exception methodology where we use these broad exceptions of medical reasons exceptions to allow for those other co-occurring conditions that could justify the prescription of an antimicrobial, say an existing infection.

Or I know Dr. Rosenfeld, if you're

looking a specific examples, but we utilized, again, that kind of broad exception category to allow for individual clinician judgment.

MEMBER GOLDBERG: So, you're stating any other diagnosis in addition to otitis media with effusion?

MS. KAYE: It's what it allows for. I can give some examples of what was from out testing data if that would be helpful.

MEMBER LYNCH: Yes, please.

MS. JIN: So, we went back and took a look at the data sample and then made a determination for the exception rate which was 11.43 percent for Measure 0657.

And there were a couple of different types of reasons for the exceptions. Specifically, individuals reported acute otitis media SNAP script, SNAP Omnicef and then sinusitis as well.

MEMBER LYNCH: Why not use those specifically instead of just medical reasons?

DR. ROSENFELD: I think there's a host

of reconcurrent things that could drive an antibiotic. So a child has a concurrent streptococcal pharyngitis, they have a concurrent bacterial sinusitis, they have a concurrent skin problem or cellulitis.

There's a lot of other reasons that kids would get antibiotics. I think trying to come up with a completely inclusive list would be very onerous.

I think the general point is that the antibiotic is being prescribe for a reason other than the otitis media with effusion.

MEMBER LYNCH: Okay.

CO-CHAIR MERENSTEIN: Go ahead, Andrew.

MEMBER SCHACHAT: Does the CPT Code that says I didn't prescribe an antibiotic just say that period or does it say I didn't prescribe an antibiotic for otitis media, in which case that solve that.

MS. KAYE: I think I can call on Kendra who is our specifications expert for this particular coding question.

MS. HANLEY: The code is general but it's intended to be used in conjunction with this measure. So, it says systemic antimicrobials not prescribed. It would need to be linked, yes.

MEMBER STEWART: It does seem that this could be improved by in not stating it not as for medical reasons but the antibiotics were -- the exclusion is not just for quote, medical reasons, the exclusion is to treat another problem or something along those lines which would be a clearer -- it would give people less leeway to say, well, in my medical opinion, I thought this patient did, if there's another problem needing antibiotics.

MS. KAYE: One thing that we can consider, and many of our measures we do, we provide examples to kind of clarify the intent of the exception with the understanding that those examples would not be an exhaustive list of all possible medical reasons but rather just to give a flavor of potential medical reasons.

Because, like Dr. Rosenfeld mentioned,

I think it's a little onerous to try to create an exhaustive list of all of the reasons.

MEMBER STEWART: I wasn't suggesting an exhaustive list, I was suggesting the different terminology instead of just for medical reasons, like if the exception was to treat another condition or something like that. Make it an active instead of a -- not to list all of the conditions.

CO-CHAIR MERENSTEIN: Any further questions?

So, we should vote on validity.

CO-CHAIR MERENSTEIN: Yes, one thing before we vote.

DR. ROSENFELD: So, in the data that I mentioned that we have on the ambulatory care visits, and this may help address the issue that was raised.

The prescribing was 32 percent of all visits, of OME visits. They also looked at prescribing for non-EMO visits, so all the other times a child showed up and that was 13 percent.

So, we could potentially anticipate that as much as 13 percent of the time, there any be a reason for giving another antibiotic, not necessarily because those children are showing up for that reason. The other ones are showing up for OME.

But even if we look at the -- take that out, we've still got a gap of about 20 percent at the worst case scenario that's being strictly given for the OME.

CO-CHAIR MERENSTEIN: Can we vote now please?

MS. ROBINSON-ECTOR: Okay. Voting for validity for Measure 0657 is open.

And for those on the line, option one is high, two is moderate, three is low and four is insufficient.

All the votes are in and voting is now closed.

Twenty percent voted high, 47 percent voted moderate, 33 percent voted low and zero voted insufficient.

So, for the criterion validity, Measure 0657 passes.

CO-CHAIR MERENSTEIN: So, Michael, we're going to go on to feasibility.

MEMBER STEWART: It seems that these data elements required are generated during delivery. It seems that this would be feasible measure to implement in practice based on what we heard, based on the group discussion we had on our call.

CO-CHAIR MERENSTEIN: Seth, anything to add?

MEMBER GOLDBERG: I agree. They've been paper-based in the past but they are easily transferrable to electronic medical records.

MS. KAYE: And I would just add that building on, you know, when these were in the PQRS program, there are existing claims specifications for this measure. It's just that because of to match our testing data, we weren't able to include those but those do exist and that helps.

CO-CHAIR MERENSTEIN: Any further questions about this?

MS. ROBINSON-ECTOR: So, voting for feasibility for Measure 0657 is now open.

And for those on the line, option one is high, two is moderate, three is low and four is insufficient.

All the votes are in and voting is now closed.

Thirty-three percent voted high, 27 percent voted moderate, 33 percent voted low and seven percent voted insufficient.

That's grey zone.

DR. WINKLER: Yes, that falls into a consensus not reached. This is not a mandatory having to pass to move on. But there does seem to be lack of consensus in the feasibility of this measure among the committee.

Anybody want to comment on that?

MEMBER SCHACHAT: I may have misinterpreted the question, but since it's a paper measure, I didn't understand how you could

use electronic sources.

DR. WINKLER: You know, I mean I think there is the what they tested, that's really the data, hard data that you have versus personal experience, the sense of how things could potentially be implemented can factor into your thinking.

MEMBER SCHACHAT: So, if we're allowed to look forward to whether it could be implemented electronically, I'm happy to vote in favor of it. But I thought the question I was having was the measure as written.

CO-CHAIR MERENSTEIN: Yes, I think it is.

MS. KAYE: If I can speak to, you know, in terms of the data, the elements required for the measure, we certainly agree that I think it could be implemented electronically.

And in terms of paper medical records, just kind of to highlight that the NQF criteria themselves don't favor one modality over another in terms of, you know, because the testing data

was paper isn't necessarily a strike against it for feasibility as long as it could also be captured electronically.

DR. WINKLER: Do they want to revote it?

MEMBER STEWART: Yes, I'm sorry, I thought that, at least what we said earlier, which is that, yes, it was done through paper, but what we were voting on is today, is this a feasible measure?

DR. WINKLER: Well, if someone -- I think, think of it this way, is the concept of feasibility is how easily is it to implement this measure? What are the burdens associated with implementation?

What you have in specifications that would allow you to implement it with chart review has been the way that they've tested it. So, that's what we know from experience and we have hard data. All right?

So, the question to you is, how feasible is this measure to be used?

MEMBER SCHACHAT: I just think chart review and using paper is generally not feasible. But I can't believe people going forward would do that. I'm sure that people will do this electronically. So, I'd be happy to support it if we could be forward looking.

MEMBER STEIN: So, this is an example where I think we're trying to tell the developer to do it a different way but we may not want to vote down the whole measure.

MEMBER FRIEDMAN: So, we probably all agree that electronically, this is easy to implement. But the way it is right now, with paper, it's difficult.

So, based on that, that's what we're asking, is do we vote feasibility favorable or unfavorable?

DR. WINKLER: I think the fact that you're not able to say totally one way or the other reflects that what you just said is what we know about the measure and the limited experience with the measure is being done, you know, through

the paper, through the specifications as is.

We don't have experience with the measure, although it seems likely. We don't have that experience.

So, the fact that you're not able to really feel comfortable one way or the other probably reflects the reality of it.

MEMBER STEWART: I would disagree that we are uncomfortable with the reality. The data for this was measured in 2004. So, we were forced to use paper records.

I think today, we're voting on the feasibility of this measure going forward. And I don't think there's people thinking, I don't know if this is feasible to be measured in today's world. I don't think anybody has any doubt about that.

We're not -- this is an old -- data for this are relatively old, EMRs were not widely in use at that time. So, I don't think that the spread of votes is because there's concern about the feasibility going forward. I think the

spread of votes is about the concern that data records were used, paper records were used for 2004 which I'm not sure is the pertinent question today.

MEMBER GOLDBERG: And just to give a concrete example, I don't see that it would be any problem with the developer going to an organization like Kaiser and saying, could you test this for the next year? And they could easily do it because they are electronic.

DR. ROSENFELD: I think we can make an analogy between this and the measure we just voted on for acute otitis externa which was basically a similar concept, don't give the oral antibiotic.

Both acute otitis externa and otitis media with effusion have a very specific CPT Code, it's 381.10 for otitis media with effusion. It's an unambiguous CPT Code.

We have the existing CPT 2 Code for not prescribing the antibiotic. So, from just a conceptual sense going forward, if there's a

desire to come up with some specs and get this into an EHR or a health care system, it would seem to be very straightforward. Famous last words, but it's basically completely analogous to the acute otitis externa, I think, experience.

MEMBER STEIN: So, there is like a post-call period and then we're going to have an opportunity to revote? Is this something that the developer can go and test electronically and then report back to us in a month or two?

MS. KAYE: I mean with testing projects like that, they're very time and resource intensive to plan, implement, analyze, you know. We, unfortunately, are measure developers but not implementers. So, we rely on external partners to conduct these testing projects as we don't have access to electronic data ourselves as stewards.

And so, you know, there's a lot to it rather than -- I don't think that's something we could certainly do before the, you know, the post-comment call. Those are pretty -- it's a

pretty intensive process.

MEMBER MADONNA: So, based on that, what you're saying is it may be quite a long time before the data source changes to electronic record? So, should we then be voting on what's going to happen in 2019 or are we going to be voting on what's in front of us right now?

MS. KAYE: I mean, one exception to that would be if, you know, if there's a registry out there or a system would chose to implement it electronically and then make their data available to us.

And so, but for us to have our own accord to go and kind of start the testing project from our end is a little bit of a different -- you know, to try to get the measure implemented versus an implementer choosing to implement that measure and then providing the data that we could then test.

MEMBER CARNAHAN: So, is this more of a question of is it feasible or do we believe it's feasible with an EMR just because the study

has a paper and we're saying it's very similar to the prior measure so it should be very feasible.

What we don't know is, is it usable? Because maybe there's no PQRS or someone to take this extracted data. That's not how it should be done.

MEMBER FRIEDMAN: And if you look at the -- point of clarification -- if you look at the verbiage here, it says under feasibility, could be captured without undue burden.

So, if we believe that electronically it could be captured, then we vote favorably for this.

MEMBER STEWART: I think that's the whole thing, could this be captured? I don't think we need to ask them to go back and do a study to see if it could be captured. Because if your ICD-9, this is all extracted from the record, it could be done.

The data is old, it was done with paper. We're not saying it must be done with paper. That's all they had.

CO-CHAIR MERENSTEIN: So, can we revote with a note that we're voting on the ability for them to capture electronically?

MS. ROBINSON-ECTOR: Okay, one moment, please. Let me clear the --

MS. LUONG: So, we will be revoting on feasibility.

And for those on the phone, if your votes for feasibility remains the same, please just let me know and if not, then please email me your new votes.

Vaishali -- okay, great.

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

Thirteen percent voted high, 60 percent voted moderate, 27 percent voted low and zero voted insufficient.

So, for the criterion of feasibility, Measure 0657 passes.

CO-CHAIR MERENSTEIN: We'll do usability.

MEMBER STEWART: So, I guess this one

is tough because it's not being publically reported now, so we don't have any data on which to base this.

It's a very prevalent condition with a clear diagnostic criteria seen frequently in primary care specialist offices.

I guess, certainly on the call, the group felt this was usable. But we didn't have data to support that.

So, I guess the question is, what are we voting on today, what the data tells us or do we think that this is a usable measure going forward?

DR. WINKLER: The use and usability criteria includes several things. The question is, how is it currently being used? You know, is it being used in programs? Is it being used for QI? Is it being publically reported? So, that's one element.

The other element of use and usability is what have we learned from the use of the measure over time. Have we seen trends? Have we

seen -- or what might be the impact of the measure? And the other thing would be, do we know of any unintended consequences?

So, use and usability covers those variety of issues that can impact the usability of the measure.

MS. KAYE: And, again, I guess I might just reiterate that while they currently are not in PQRS, again, because it's a little tougher to find a home for pediatric measures when they aren't captured in that Medicare population.

And so there are currently efforts underway to get them in the Medicaid programs with states, to try to get these used in the, hopefully, near term future. So there are some things underway, we just, unfortunately, don't have anything concrete at this moment.

CO-CHAIR YAREMCHUK: So, I have a question. There are also HEDIS measures, and that applies to the Medicaid population. So, has this been used in the Medicaid population for NCQA?

MS. KAYE: Not that I am aware of. But that would be something that maybe I would need to confirm.

CO-CHAIR YAREMCHUK: Okay.

MS. KAYE: But off the top of my head, I'm not sure.

CO-CHAIR YAREMCHUK: Okay. And then my other question is to Reva. And this is kind of a Committee question. But how often are measures going to come forward to this Committee? Is this going to be an annual type? Is it once every four years?

DR. WINKLER: The typical review is three years. If something were to come up in the intervening timeframe, a material change to the measure, including additional data sources or someone externally requests an ad hoc review because of an issue that has arisen, unintended consequences, new evidence, something out there, then we might do it sooner.

MS. KAYE: And I guess to add a little bit to your HEDIS question, the NCQA, who kind of

governs, they do a lot of their own measure development. And so, for that program, they usually do homegrown measures rather than incorporating those developed by other.

CO-CHAIR YAREMCHUK: Reva may be able to speak to the -- many of the HEDIS measures are NQF and there's a harmonization of some of the measures.

DR. WINKLER: Certainly. And you're going to see one in a little bit. You know, NCQA brings a lot of their HEDIS measures to NQF for endorsement.

But we see measures developed by a wide variety of developers who have, you know, a particular niche, oftentimes, and different relationships with how measures get implemented. And we don't have a -- we don't limit that. And so we do see measures coming in that do have different characteristics.

MEMBER CARNAHAN: So, the previous measures all had a legacy to them and that made it easier for us to know they'd be implemented.

And so for newer measures where they don't really have a set program, you can't really get that going unless you have a measure. So, how do you balance that when we're looking at usability?

DR. WINKLER: I think that there's the issue between new measures, which I think those questions are quite valid because we wouldn't have much of a track record, versus measures that have been endorsed and have been available for a while. I suppose one of the questions might be, you know, what are the limitations on it being picked up for use? Why don't we have much of a track record? I think they've spoken to a certain amount of that.

And the question of, is there potentially a path going forward that can change that? Can we anticipate things being different down the road? And, again, I think they've spoken to that as well.

MEMBER STEWART: I would say, certainly, that, you know, going forward this

seems quite usable. It's just that we don't have a track record because it's a pediatric condition and Medicare didn't pick it up in PQRS. And so, you know, we can't base it on the data.

But certainly it seems like, if endorsed, this measure would be usable and useful, because, in fact, it is a prevalent condition where treatments are frequently over prescribed that have harm and costs.

CO-CHAIR MERENSTEIN: We can vote on it, unless there's other questions or comments.

MS. ROBINSON-ECTOR: Thank you. Voting for usability and use is now open for Measure 0657.

And for those on the phone: option one is high, two is moderate, three is low and four is insufficient information.

MS. LUONG: Vaishali, if you can email me your votes, thanks.

MS. ROBINSON-ECTOR: All votes are in and voting is now closed.

Seven percent voted high, 80 percent

voted moderate, seven percent voted low and seven percent voted insufficient information.

For the criterion of usability and use, the measure passes.

CO-CHAIR MERENSTEIN: Any discussions before we vote on the last one?

MS. ROBINSON-ECTOR: Voting is open for overall suitability for endorsement for Measure 0657.

And for those on the phone, option one is yes and option two is no.

All the votes are in and voting is now closed. Eighty-seven percent voted yes, 13 percent voted no.

So, for recommendation for overall suitability for endorsement for Measure 0657, the measure passes.

CO-CHAIR MERENSTEIN: Let's go on to the next one before lunch.

We're going to do Otitis Media with Effusion, systemic corticosteroids. And feel free to ask questions, but I think we've talked a

lot about it. Maybe the usability and feasibility we can agree on from before, but we'll see what happens. Todd?

MEMBER RAMBASEK: Should I shorten that part of the discussion?

CO-CHAIR MERENSTEIN: I think you should. I mean, it's up to you, but I think we've talked about it.

MEMBER RAMBASEK: I agree with you wholeheartedly.

So, this is a process-based measure suggesting that we limit the use of systemic steroids for otitis media with effusion.

The evidence is strong, based on ten randomized control trials published in Pediatrics in 2004. So I think there is little question that there is strong evidence based on the fact that systemic steroids provide little benefit and significant harms. So, I think we can vote on that.

CO-CHAIR MERENSTEIN: Anything to add, Kathy?

CO-CHAIR YAREMCHUK: The other thing that we've talked about is -- and I'm looking at what the 2004 guidelines and the 2011 Cochrane Review -- and they lumped together antimicrobials and corticosteroids. And so I think it's been raised before whether we need two different measures, or one measure and lump those two together.

DR. ROSENFELD: Just to respond to that, in the 2004 guideline, I think primarily due to lack of experience -- and people weren't thinking about performance measures in 2004 -- actually, probably in 2002 when that process started it was more just in the conceptual thought of, yeah, let's condemn a whole bunch of medicines in one statement.

Since then, the evidence is a little more robust and the importance of having actionable guideline statements that can be a basis for a valid measure has become paramount.

So, that's why it's split into three separate ones. And, as I mentioned, the

currently unpublished update of that guideline, to be published probably in January or February of 2016, will have three separate strong recommendations, individually, for the different medications.

CO-CHAIR YAREMCHUK: And I guess that's just my question when I had asked Reva before about how often we're going to do this, because it would be nice if we had the 2016 now that had the three different recommendations and we could probably get different information about it.

CO-CHAIR MERENSTEIN: All right, we're going to vote now on the evidence.

MS. ROBINSON-ECTOR: So, voting for evidence is now open. And for those on the line: option one is high, two is moderate, three is low and four is insufficient.

All the votes are in and voting is now closed.

Sixty percent voted high, 33 percent voted moderate, seven percent voted low and zero

voted insufficient.

So, for the criterion of evidence, the measure passes.

CO-CHAIR MERENSTEIN: So, we're going to go opportunity for improvement.

MEMBER RAMBASEK: So, I feel the performance gap is the most difficult part here that will require more discussion.

So, we have data presented by the developers that show, in 2008, in Otolaryngology-- Head and Neck Surgery, we have a ten percent use rate for steroids. So, that's one data point, 10 percent in 2008. And then Dr. Rosenfeld has helpfully shared more data with us that we're now down to 3.2 percent.

So, my understanding of that database is that, recently, looking at over 100,000 OME visits, that 3.2 percent of those patients got steroids.

And, as I understand it, probably some perhaps small fraction of those patients had a condition for which they may have needed

steroids. So, maybe somebody with otitis media with effusion also has asthma that day and had a valid reason for steroids.

So, that 3.2 percent that got steroids, maybe if we say a small fraction actually needed them, maybe 2.9 of patients with otitis media with effusion were inappropriately given systemic steroids.

So, I think that really calls into question whether we want to expend the energy to chase that two to three percent.

CO-CHAIR MERENSTEIN: Rich?

DR. ROSENFELD: I agree completely that the absolute number is somewhat depressing and not probably enough to just on itself sway this erudite body. But I think, in this condition, it has to be looked at a little differently.

If we, again, look at the prevalence of this condition, which is somewhere between 50 and 90 percent of children get this in their early childhood. We have 25 million kids currently under five -- is it under five, let me

get the number right -- no, that's -- yes, under five, 25 million children under age five in the U.S.

So, even if we go with 50 percent, that's about 12 million cases a year. If we take three percent and apply it to that, it's about 360,000 children a year potentially getting systemic steroids, a very minute portion of which, or maybe not so minute, could suffer things later in life, such as cataracts, immune suppression, aseptic necrosis of the femoral head.

And, again, to me, the most of concern is the disseminated varicella, chicken pox, which does kill kids and there are ample reports of that. And there is this ridiculous anti-vaccination trend going on which makes that ever potentially more of a concern.

So, I think the number is small, but the condition is highly prevalent. And the downsides of even a small amount of systemic steroid prescribing for this have the potential

to be very, very serious in children.

I don't know how you balance all that in thinking about this, but, again, to me, the real benefit of having a little benefit of the doubt here and getting this measure through is that it will encourage dissemination of knowledge, that this is a bad idea, and we will, down the road, have data as this gets picked up and used.

And in three years, if we find out it's a non-issue, so be it. The measure disappears. But we may never know the answer to this question of what's the variability among practitioners unless we put some incentive out there for people to actually tell us what they're doing.

CO-CHAIR MERENSTEIN: So, I guess I would disagree. I would agree with Todd that no one's arguing that it's a bad idea to do it, it's just you have this limited amount of resources. It seems like to me we should be going over against antibiotic overuse, which is much more

prevalent, is a much more serious issue.

Three percent, to me, is almost a hundred percent. I mean, you can't get any better than that. Now, that's not published data; the published data we have is ten percent, right?

But three percent, no matter what NQF does, I don't know how you get better than three percent of usage, really.

MEMBER MADONNA: The other thing that I would add is that we appear to have a trend, if we take those two data points, and the trend drives with the story and the evidence.

There were some initial studies that suggest a little benefit. People started doing it. And then more studies came out saying maybe there's an initial small benefit, but no long term benefit. So the practice patterns have been changing independent of us, independent of what we do.

So, CME and medical education is already working, it's already disseminating the

information, and we might not need to push harder on something that's on the needle that's already moving in the right direction.

CO-CHAIR YAREMCHUK: I agree with what Todd and what Dan has said about this.

CO-CHAIR MERENSTEIN: Steve?

MEMBER PATEL: So, this is Vaishali on the phone. I have a question, actually.

How sure are we, you know, that three percent, that number from the claims database, actually is an accurate number that's associated with, you know, use of steroids in this condition? Could it be from other things, too? Because if that's possible, then even the three percent is too high.

DR. ROSENFELD: There is some limited data on that. Part of the analysis that was done was also to look at the rate of systemic steroid prescribing in all of the non-otitis media with effusion visits, which was about half that. It was about 1.7 percent.

So, at the very least, there's a gap

of about 1.5 percent. It's probably much higher than that because the 1.7 percent receiving steroids for their primary diagnosis are probably not the kids with OME. They're kids showing up with croup. They're showing up with asthma attacks and other reasons for steroids.

The OME one of 3.2 percent is with a primary diagnosis of OME. So, some of it may be due to other conditions. If we get the most generous possible, it's about 1.5 or 1.7 percent. But it's probably a small number.

MEMBER PATEL: Yeah. Okay, thank you.

CO-CHAIR MERENSTEIN: Steve?

MEMBER STRODE: Yeah, continuing on that, the ten percent figure from 2008 was otolaryngologists. The more recent data where it gets down to three percent, 1.7 percent, is that still a population of otolaryngologists or does that include primary care pediatricians and family doctors?

DR. ROSENFELD: Let's see. So, it includes -- the only thing they looked at

initially on these data are the otolaryngologists versus the emergency physicians. And the otolaryngologists on this were about 2.5 percent. Oh, we have some more data. The non-otolaryngology visits were 3.6 percent. For just strictly the non-otolaryngologists. So a little higher if we look at that group.

MEMBER STRODE: But would that be a non-ENT population that was just emergency physicians?

DR. ROSENFELD: No, this database had millions of encounters in it. 114,000 of those were otitis media with effusion. A small portion of that was otolaryngologists. So, this would include everybody else, the 3.6 percent, the non-otolaryngologists.

CO-CHAIR YAREMCHUK: Do you know if that was the only ICD-9 that was reported for that, or were there other associated? Because the question comes back, were the kids also with asthma and did they have other issues?

And so if the only ICD-9 was otitis

media with effusion versus primary diagnosis of OME, the second one was asthma exacerbation, the third one was something else, does it muddy the waters?

DR. ROSENFELD: I don't think -- the analysis hasn't been done at this point to look at the prevalence of the concurrent conditions.

CO-CHAIR MERENSTEIN: Josh?

MEMBER STEIN: It seems like these measures are mostly being done by ENT docs. I mean, I know you couldn't answer that question or the earlier one. So, if the ones that are misprescribing the steroids are not the ENT docs, are we really capturing the quality and the intended audience? Do you guys want to comment on that?

DR. ROSENFELD: I think what's been done in the past is, as we've discussed, extremely limited as far as the paper-based. But this condition is a mainstay of pediatric practice and family practice as well.

You know, I understand maybe about 20,

30 percent of family physician visits are kids, and probably the most common thing among those, other than the typical viral upper respiratory infection, would be some form of otitis media.

So, it's ubiquitous. It's seen in the emergency departments, as well. I think probably a minority of otitis media with effusion care is provided by otolaryngologists in real world practice. It's even seen by audiologists as well.

MEMBER YOUDE: Yeah, I have a couple of comments. First is a clarifying question just to make sure I'm understanding it. So, that ten percent that we originally looked at was ENTs alone and then that three percent that we looked at was ENTs plus other fields.

I would actually be more interested in the patient number associated with those percentages. Because while we're seeing ten go down to three, we're also changing what we're including.

And so, if that's possible, to know

the actual patient numbers associated with those percentages or what the grouping is there, that would be interesting because we may actually see a reverse trend.

Also, I know that we have limited resources and just in my practice alone, obviously, I'm not an ENT, I'm an audiologist, but I have seen children prescribed with, you know, steroids and whatnot.

And so, as an audiologist, knowing that that's not right, I'm sitting here looking at who is my patient and what are they going through? And then I look at it as a mother and I'm like, okay, if my kid's getting steroids this young, I would be furious.

And so I know that we're looking possibly at a small percentage, and I know that we're looking at limited resources, but I can't help but -- I don't want us to lose sight of who this is and who this impacts given the consequences that are associated with it.

CO-CHAIR MERENSTEIN: I would agree

with you that viral pharyngitis, I think, just from anecdotal evidence, is you would prescribe steroids a lot more than this. If we could do three percent, I'd be happy with that for viral -- I think it's much, much higher.

So, I think there's better numbers to decrease -- I'm not arguing that we should recommend steroid use for these things that are overused. I just think that the numbers are relatively low for this.

MEMBER PATEL: This is Vaishali on the phone. I would also agree with that.

CO-CHAIR MERENSTEIN: Can we vote on it, then?

MS. ROBINSON-ECTOR: Voting for performance gap for Measure 0656 is now open. And for those on the phone: option one is high, two is moderate, three is low and four is insufficient.

All the votes are in and voting is now closed.

Thirteen percent voted high, 20

percent voted moderate, 40 percent voted low and 27 percent voted insufficient.

So, for the criterion of performance gap, the measure does not pass.

DR. WINKLER: Given the conversation, there's another twist for this. This is an endorsed measure. We do see measures being somewhat successful out in the field as a result of either change in practice, whatever.

But if the measure is still a good measure, we do have the opportunity of giving it a tag of inactive endorsement such that it stays on a list of endorsed measures, but we're sending the signal that, hey, the opportunity for improvement here may not be all that great and you really want to be careful about implementing something that could not have a great cost-benefit ratio.

So, any endorsed measure that fails opportunity for improvement is then eligible to go into this inactive endorsement status. And so that would mean, you know, being sure that it

passes all the rest of the criteria, and then you could recommend it for that inactive endorsement status.

CO-CHAIR MERENSTEIN: Is that a question? Or are you saying we should go on?

DR. WINKLER: Well, does anybody have any comments or questions before I move on? I said a lot of stuff. Okay.

So, we want to go through the remainder of the questions and criteria, and then you're final question would be, do you want to recommend it for inactive endorsement?

CO-CHAIR MERENSTEIN: So, we'll do reliability.

MEMBER RAMBASEK: Okay. So, the denominator here is patients two months to 12 years with otitis media with effusion and the numerator is the number not treated with steroids. And the exclusions were generally the coexistence of another steroid requiring condition.

The reliability was measured amongst

two large pediatric networks. And, in doing that, there were 114 re-abstractions performed, and there's a high rate of correlation between the re-abstractions. When one person abstracted a chart, when a second reviewer re-abstracted that chart, they almost always got the same answer as indicated by the kappa statistic of 0.85.

So, I think the reliability is high, and unless there's questions, we can vote on that.

CO-CHAIR MERENSTEIN: Kathy, anything to add? Any questions? We can vote.

MS. ROBINSON-ECTOR: Voting for reliability for Measure 0656 is open. And for those on the line: option one is high, two is moderate, three is low, and four is insufficient.

All the votes are in, and voting is now closed. Sixty-seven percent voted high, 33 percent voted moderate, zero voted low and zero voted insufficient.

So, for the criterion reliability, the

measure passes.

CO-CHAIR MERENSTEIN: Validity?

MEMBER RAMBASEK: So, there are a large number of studies suggesting to us that this is clearly aligned with the evidence. This was assessed primarily using face validity by asking members of the American Academy of Otolaryngology Patient Safety Committee. And they basically gave it a very high rating. It had very good face validity to members of that committee that, if this indeed was done, if systemic steroids were avoided, that that probably had a high correlation with good quality care.

So, there were very few exclusions with regard to threats to validity. Really, the only main exclusion being coexistence of another steroid-requiring condition. And there was no risk adjustment performed because I don't think that's really appropriate for this sort of measure. Kathy?

CO-CHAIR YAREMCHUK: Nothing.

CO-CHAIR MERENSTEIN: Any questions about this before we vote? Okay, we'll vote.

MS. ROBINSON-ECTOR: Voting for validity for Measure 0656 is now open. And for those on the phone, option one is high, two is moderate, three is low and four is insufficient.

All the votes are in. Forty percent voted high, 60 percent voted moderate, zero voted low and zero voted insufficient. So, for validity of Measure 0656, the measure passes.

CO-CHAIR MERENSTEIN: So the question was, should we just use the same votes for feasibility? Or do you have anything else to add about the electronic --

MEMBER RAMBASEK: No, nothing else to add. We can use the same votes, that's fine.

CO-CHAIR MERENSTEIN: Same? Do you want to vote again, or do you just want to use the same vote? Just use the same vote? Everyone okay with that? And usability. It's really the same issues for the next one, too.

MEMBER RAMBASEK: Yes.

CO-CHAIR MERENSTEIN: All right, so we'll go to the last vote.

MEMBER LYNCH: I'm sorry, what are we voting for now?

CO-CHAIR MERENSTEIN: So we're actually not voting yet. We're going to vote for the overall suitability for endorsement.

DR. WINKLER: And just with the caveat that this would be an inactive endorsement status. It's still endorsed, but the inactive status indicates your concern about the opportunity for improvement is high, and so it should be used with caution.

CO-CHAIR MERENSTEIN: There's a question.

DR. RICH: I think you're actually touching on something that's really important --

CO-CHAIR MERENSTEIN: I'm sorry, you need to introduce yourself.

DR. RICH: I'm sorry. My name is Bill Rich, Medical Director of Health Policy, and I'll be meeting you all after lunch.

But if you have a really good measure, even if there's not an area for improvement, there's really robust evidence that says that if you drop the measure the performance immediately falls. There's an article in British Medical Journal where Joe Selby from Kaiser, when he was at Kaiser, took the Kaiser database and the UK database and showed that once a measure was removed from P for P, the performance immediately falls at about three percent a year.

Unpublished data from Kaiser in Colorado, they had the best outcomes in the United States for control of hypertension. Once they hit that, they dropped it, their control fell. So I think that's one thing to be aware of if you have a good measure, like Dr. Winkler pointed out, there are some consequences if it's no longer active.

CO-CHAIR MERENSTEIN: Any questions or comments? Tammy?

MEMBER BRADHAM: Depending on how the vote goes, can the group come back and resubmit?

DR. WINKLER: Sure, we can always reactivate it, depending if there's new information or something to work on, new data. Again, absolutely. This is very dynamic, it's been a process that comes and goes. As new information becomes available, certainly.

That's one of the benefits of having you all constitute as a standing committee: if something comes up, we can come back and bring it back in.

Also the other thing is, with your recommendations, we may get feedback from public comment that you may want to factor into your final decision. When we regroup after the call, there may be public comment that may influence how your recommendations go.

CO-CHAIR MERENSTEIN: Josh first.

MEMBER PATEL: So before we rush through and finalize this -- this is Vaishali on the phone -- you know, I have to say I was a little bit confused about us being out of sequence earlier for voting on this measure

overall as opposed to going one step at a time, reliability, validity, et cetera. So I would request, if it is possible, to do a re-vote on this measure.

And then, again, the other comment is I agree -- I think that was Bill Rich -- and I agree with his comment wholeheartedly that, you know, if we have a good measure which is actually giving us outcomes, removing that measure may actually have things go in the reverse direction. So that is something to be aware of.

And my other comment is that, yes, we can get -- I think it's in line with Bill's comment, again -- is we can look at claims databases and we can estimate rates of utilization of steroids, but the one big limitation with that is that steroids, if they're oral, they're prescribed and they appear in pharmacy claims data, and the actual diagnoses codes don't appear in pharmacy claims data; they appear in medical claims data, okay?

So, at best, you can only draw

correlations of what it might be used for, but you don't know 100 percent. And claims databases, unless you have evidence from 10 different claims data that all point to the same thing, it's really risky to make a decision based on one claims database study or two claims database studies that the utilization, with certainty, is only three percent, or only 10 percent. And this is coming from somebody who does a lot of claims database research.

CO-CHAIR MERENSTEIN: So, we have two comments. I would just say that we have voted, and it really wasn't that close, that one part. So unless there's someone that would second that, I don't see any reason to re-vote that issue.

DR. WILSON: Could I just suggest that we clarify what we did just vote on for this specific measure, for the transcript and for the record. This is what we've done, not going back through the votes, but we have voted on importance, we have voted on scientific acceptability. And my understanding is that the

group elected to, for feasibility and usability, vote as they did on the last measure. Have I got that clear?

CO-CHAIR MERENSTEIN: Yes, which means it passed. Yeah.

DR. WILSON: Correct. I just wanted to clarify, because it sounded like there was some confusion about what we had voted on for this particular measure. Thank you.

CO-CHAIR MERENSTEIN: All right. Josh, then Andrew.

MEMBER STEIN: I just wanted to clarify the process. If the developer wants to tweak the measure, can it be re-assessed on the post-call so it may not get to a point where it's taken out of practice?

DR. WINKLER: Yes, there can be responses during the whole process where you have the opportunity to change your recommendation.

CO-CHAIR MERENSTEIN: Andrew?

MEMBER SCHACHAT: I wanted to understand the implication of putting it in

reserve status. Does that mean that Medicare is not likely to use it, or -- so, my anxiety is that, even though there's no opportunity for improvement, the argument makes sense to me that if you stop measuring it, that performance is going to go down. Why not keep things?

DR. WINKLER: That's definitely the discussion for you all to determine.

MEMBER PATEL: So, again, this is Vaishali, and I'm confused. So have we voted on the actual, you know, the measure approval? Yes or no?

CO-CHAIR MERENSTEIN: No, we haven't voted on that yet. We are going to vote on that.

MEMBER PATEL: Okay, so why was there comment regarding the fact that we're voting on us, you know, liking the measure only, not really keeping the measure?

MS. LUONG: So, with the measure, we voted on the first criteria, which includes two sub-criterion under it. Evidence passed, while performance gap did not, it was voted down. And

so the whole discussion right now, Vaishali, was that Reva brought up an option for the measure to still go forward, and we can vote for it as reserve status. And Reva, you can speak a little bit more to the reserve status.

MEMBER STEWART: Can I ask a question too about -- and it follows on what Andy was just asking. I think Dan brought up the point, well, we have limited resources, so why should we support something where we've already achieved the goal? But I guess the question is, what resources would we use if we kept it as an active measure? I mean, are we really using resources? Because it's already an active measure, is there any negative consequence to us continuing to endorse it?

DR. WINKLER: The resources we're talking about, the cost-benefit balance, is about the cost of collecting data, analyzing data, and reporting data. And so there is a cost associated with actual measurement. And so, you know, all measurement has costs.

MEMBER SCHACHAT: But those are optional costs. Anyone who doesn't want to do the measure doesn't have to do the measure, and they avoid the costs, I guess.

I just think that -- I buy the argument that if you stop measuring things, performance goes down, and this doesn't seem to be all that difficult to allow it to continue.

So what I'm putting on the table is your rule set about if the performance gap is small, you shouldn't endorse things. And I'm not sure that makes sense, especially when there's a large --

MEMBER PATEL: I agree.

MEMBER SCHACHAT: -- when there's a large n. It's a very common disease.

DR. WINKLER: That's why it's a discussion point for you all, but one of the criteria is around opportunity for improvement, and so you need to weight that. There are no absolutes. So that's one of the questions for you all to determine.

MEMBER MADONNA: Isn't that just a philosophical thing for every measure that we look at? I mean, ultimately, this is about improving care, so what we really want to do is get everything up to, you know, theoretically, 100 percent. If Bill is correct, and I'm sure he is, that is a measure is dropped, performance goes down, then, you know, you're looking at thousands of children who are going to be over-treated with corticosteroids.

So I think this is really -- going back to the first thing I said, and piggybacking off of Andy, really this is philosophical in terms of every measure.

CO-CHAIR MERENSTEIN: Josh first. Todd, I'm sorry.

MEMBER RAMBASEK: That's okay. So I think that's a key point: when you stop measuring something, does it regress? But this is not publicly reported now, is my understanding, so I see this as being used in the ABIM maintenance and certification, but I haven't seen it as

publicly reported anywhere. So is it really something that is already pushing on practice patterns, in its current incarnation, in the way it's been -- in other words, did the 10 percent go to three percent because of this? Or because of general dissemination of information? And I think that would require someone in the field of care to answer that.

MEMBER FRIEDMAN: So, I think we're going back and forth. So, if people want to take this measure and vote whether it should be active still, then we probably have to go back and look at opportunity for improvement again and re-vote. And if we think that we voted incorrectly on that, we can go back and improve it, and then we can move forward with active endorsement.

Right now, the way it stands -- and Reva you can clarify it -- opportunity for improvement was voted down. So at this point, we can't vote for an active endorsement. We can either put it in reserve, or just nix it at all.

DR. WINKLER: You said it better than

I could.

MEMBER YOUDE: Yeah, I do have a comment. So when I'm looking at this process metric -- and I look at process metrics an awful lot -- I think of process metrics as, where are we at in holding the gains? And so what's going on is like, okay, so let's say that went 10 percent to three percent. That's awesome. Now let's take it away.

What happens when you take away process metrics is people start going all willy-nilly doing what they want to do, practicing how they want to practice, how they feel is fit. And so we do have the risk of falling backwards. So basically the process metric is holding the gains that we're seeing. If we're saying it's 10 to three, then we're holding at three. And from what I'm understanding in this discussion is three is pretty great.

CO-CHAIR YAREMCHUK: Okay. And I guess it's going back to what Todd had said: it's not clear that the process measure, if it's not

publicly reported, is what's driving it, or is it public knowledge?

And I appreciate that if you don't measure it, you don't get to perfection. At the same time, we use the example of immunizations. We've been measuring that forever, and part of the issue is there's been a change in culture or issues or whatever you want to describe, and that's -- the measuring of it hasn't changed, you know, the incidence of it or the prevalence of it. What's changed is an outside factor.

So, I think measurement is one process that drives improvement. I don't think it's the only process that drives change in patterns, change in performance and that kind of thing. So I think the question that was raised is, since we're not publicly reporting this, but there's an awful lot out there about this measure.

And then the other think that's kind of mixed into this is that we don't know that the steroids in the incidents that we're looking at is solely for the otitis media with effusion, or

is it there's some other comorbidities and other visits that are going in. So, much like when you see a kid that has serous otitis media, is it because of that they're getting the steroids, or was it something else at the time they were seen? So that's the only concern.

CO-CHAIR MERENSTEIN: Michael, and then Josh, then we should try to vote again, I think.

MEMBER STEWART: I was just going to respond to Todd's question. I believe, as an otolaryngologist who pays attention to this, that the success here has been because there's been an Academy-endorsed evidence-based guideline, and that people believe the guidelines, they believe the process, there's a respect for the process, and it's been disseminated and it's been discussed, and people look at it and say, oh, wow, the guidelines says I shouldn't do this, maybe I shouldn't do it. Not because it's been publicly reported, because there was actually an evidence-based guideline that made the

recommendation. That's been the improvement.

MEMBER STEIN: Just, again, to clarify. Our voting it down as an active measure now, it can be re-voted on in the post-call period and be determined to be an active measure once the developer has fixed some of these issues, correct?

DR. WINKLER: Or if there is other feedback that makes you reconsider your position. That would be one. You know, I'm not getting a sense that the developer has something specific to do with the measure per se that would change things. It's more the fact that the data on whether this is a true quality problem, at this point in time, to justify measurement. I think that is the discussion you seem to be having.

CO-CHAIR MERENSTEIN: Matt.

MEMBER CARNAHAN: I think Todd makes a good point. I come from an organization where we have measure fatigue. And we have so many competing measures that, anything that happens, an act of Congress, and we kind of get reduction

in everything and we're trying to do everything. So you have to really pick the important ones.

And this is important, we agree to that, would the amount of value that this approved measure would make may not be as great as the improvement that an organization can make by saying to its members, "this is the best practice." And so maybe we do focus on the antibiotics and other ones where there's more area to gain and not get the measure fatigue associated with it.

CO-CHAIR MERENSTEIN: So, is there a reason -- do people want to re-vote this one, or just vote for the last one? I mean, it seems like, to me, we just go to the last and vote, unless there are people who want to re-vote. I don't think this vote is going to change that much.

MEMBER STEWART: Is it feasible to have three options? So, right now, we're either being asked, you know, inactive or no. Can we allowed to vote on keep it active, inactive, or

no?

DR. WINKLER: The way to do that would be to re-vote the opportunity for improvement, just as Scott described. So the question is, do you want to do that? Does anybody feel that they would change their vote to the -- we've got the results up there. Based on the last discussion, does anyone feel that they would change their vote?

MEMBER PATEL: I would.

MEMBER YOUDE: Okay. So I think we've had some new information come to light since we previously voted. And while we may not understand what that means at this point in time, the fact that we've had that discussion, to me, warrants a re-vote, just because there's been new information presented.

CO-CHAIR MERENSTEIN: I think that's fine. Let's just re-vote this one question here, and then we'll go to the last question.

MEMBER LYNCH: Are we voting now?

MS. ROBINSON-ECTOR: So, voting for

performance gap for Measure 0656 is now open. And for those on the line: option one is high, two is moderate, three is low, and four is insufficient.

MS. LUONG: Yes, Judith, you can send me your votes.

MS. ROBINSON-ECTOR: All the votes are in. Seven percent voted high, 27 percent voted moderate, 40 percent voted low and 27 percent voted insufficient. So, the measure does not pass performance gap.

CO-CHAIR MERENSTEIN: Do we just go to the last vote, then?

MS. ROBINSON-ECTOR: Yes.

MS. LUONG: So, for those on the phone, that's reserve status.

MEMBER LYNCH: Could you please clarify the different votes on this?

MS. ROBINSON-ECTOR: Yes. So, option one is yes, and option two is no. And voting is now open for recommendation for reserve status for Measure 0656.

MEMBER LYNCH: Thank you.

MS. ROBINSON-ECTOR: All the votes are in. Eighty-seven percent voted yes and 13 percent voted no. So, the measure does pass for recommendation for reserve status.

MS. LUONG: Operator, can you open the line now for public and member comments? And for anyone in the room, if you would like to speak, please come up to the microphone.

OPERATOR: At this time, if you would like to make a comment, please press star, then the number one.

There are no comments from the phone line.

MS. LUONG: There are no comments in the room, either. So, I'm sorry for the delay in time, but we will break now for lunch.

Okay, so, before we break for lunch, if you're still having problems accessing SharePoint, please raise your hand so we can just take a record vote of that.

CO-CHAIR MERENSTEIN: We're going to

try to restart at five after 1:00, if we can, so that we'll only be 20 minutes late.

DR. WINKLER: And also, lunch is back here in the back of the room.

MS. LUONG: Yes. So, five after one. Thank you, all.

(Whereupon, the above-entitled matter went off the record at 12:38 p.m. and resumed at 1:05 p.m.)

MS. LUONG: So everyone, I hope everyone's coming back soon. We will start off the afternoon session with selection of terms. I believe Shaconna has walked around the room and you all have selected from this jar just a paper, a little paper with either a two-year or a three-year term. For those that have not, can you please raise your hand, and I'll walk around.

And for Judith and Vaishali, I will be the proxy and do it for you right now. Judith, you have two-year, and Vaishali you have two-year as well, and we had one committee member who wasn't able to make the meeting today, John

McClay, so I'm going to do it for him as well. So John McClay has two years, too.

So now we will go around the room to disclose how many -- what your term is for the remaining of -- either two or three years, and if you object to more than two years, if you have a three year and you think that you can only do two years, please let staff here at NQF know via email, that would be fine, or pull us aside and you can discuss with us then. So I guess we can start now with the two co-chairs. So Kathy's three. Three. Dan's three, okay. Now Matt? Two? Steve? Two. Scott is two as well. Rich is three, Rich Madonna. And can you just speak into the phone? I can't see you.

MEMBER RAMBASEK: Todd, three years.

MS. LUONG: Three years? Okay great. I didn't see a three.

MEMBER STEIN: Josh, three.

MS. LUONG: Josh, three.

MEMBER STEWART: Micky, three.

MEMBER BRADHAM: Tammy, three.

MEMBER GOLDBERG: Seth, three.

MEMBER SCHACHAT: Andy, three.

MEMBER YOUDE: Jackie, two.

MS. LUONG: So that concludes the selection of terms. Now I will hand it back to the Co-Chairs to start the discussion of 0655.

CO-CHAIR MERENSTEIN: So we're going to start it again with again, another--last otitis media with effusion, this is antihistamines or decongestants. You guys already spoke, do you have anything to add? All you guys are good? All right, so we'll go straight to Kathy and Todd.

CO-CHAIR YAREMCHUK: So once again, this is I'm going to say the last one of otitis media with effusion, and it's regarding antihistamines or decongestants, and I think we've talked about in terms of the evidence, it's the same background, very common situation for population. Previously, the 2004 guidelines, as well as the Cochrane review in 2011 looking at the use of antihistamines and decongestants for

treatment of serious otitis media.

So the majority of these physicians do the watchful waiting that we've talked about, and about 14 percent of physicians in otolaryngology prescribes antihistamines or decongestants for pediatric patients. And I'm only going to add that probably the majority of these antihistamines and decongestants are over the counter, and won't be prescription in nature. Anything?

CO-CHAIR MERENSTEIN: So again on the evidence, unless there's questions about the evidence for this measure.

MS. ROBINSON-ECTOR: Evidence is open for measure 0655, and for those on the phone, option 1 is high, 2 is moderate, 3 is low, and 4 is insufficient evidence.

MS. LUONG: For those on the phone, can you please vote via email? Thanks. Just to confirm, do we have Vaishali and Judith back on line?

MEMBER LYNCH: Yes.

MEMBER PATEL: I'm listening, but I missed the discussion, sorry.

MS. LUONG: Oh okay, that's fine. You can just sit out for this, and then the discussion will continue with performance gap. Thank you.

MS. ROBINSON-ECTOR: So all the votes are in. 10 voted high, 3 voted moderate, 1 voted low, and zero voted insufficient. So for evidence for measure 0655, the measure passes.

CO-CHAIR MERENSTEIN: So for opportunity for improvement?

CO-CHAIR YAREMCHUK: When it talks about opportunity for improvement, there were -- I'm looking for the percent of providers that continued to use -- 14 percent of physicians in otolaryngology prescribed antihistamines and decongestants. So I'm going to say that that would be the opportunity in terms of decreasing that number.

CO-CHAIR MERENSTEIN: Any questions about that before we vote on it? I'm sorry,

someone has a question? Yes, comments are fine.

DR. ROSENFELD: I would just say that what's not apparent, as necessarily as obvious with the antihistamines decongestants are the adverse events that you also get that are maybe a little more obvious with steroids. And before coming here, I just did a little literature search, found a bunch of articles about adverse events, and it's related to the sedating antihistamines in particular, and the decongestants.

One article in 2004 and 2005, they looked at 1,500 emergency department visits, including three deaths, most of which occurred from overdosing, giving inappropriate doses to young kids, and a number of other articles talk about all the sedating and soporific properties of these drugs. As Kathy said, most of them, about 60 percent in the new data we have seem to be the over-the counter, and we do have a CPT-2 code that currently exists that allows tracking as to whether you prescribe or recommend these

drugs. So there is a code for it. Thanks.

CO-CHAIR MERENSTEIN: All right, we're going to vote on it now?

MS. ROBINSON-ECTOR: Voting for performance gap for measure 0655 is now open, and for those on the phone, option 1 is high, 2 is moderate, three is low and 4 is insufficient. All the votes are in. 27 percent voted high, 67 percent voted moderate, 7 percent voted low, and zero voted insufficient. So for performance gap, measure 0655 passes.

CO-CHAIR MERENSTEIN: Reliability?

CO-CHAIR YAREMCHUK: Reliability is-- this was done on paper charts, and so the data-- reliability was tested 2008, 2009, 98 percent agreement for the numerator, 95 percent for the denominator, so a high reliability in terms of testing chart extraction for this.

CO-CHAIR MERENSTEIN: I have a question. When you say chart extraction, this is just in ENT doctors, or this is in all doctors?

CO-CHAIR YAREMCHUK: It said pediatric

networks, and so I would think it could be pediatricians, it could be pediatric otolaryngologists.

MS. JIN: So for the population that was included, it includes all qualifying providers that are able to report or report information on those measures--that measure.

CO-CHAIR MERENSTEIN: But what kind of doctors are those that are qualified?

MS. JIN: There's no--we weren't able to gather specifics because we randomly sampled the charts, so we didn't take a look at the exact providers that were providing the information. But they were from the two pediatric network groups.

CO-CHAIR MERENSTEIN: Any other questions or comments about that? Then I guess open for voting.

MS. ROBINSON-ECTOR: Voting for reliability for measure 0655 is open, and for those on the call, option 1 is high, 2 is moderate, 3 is low, and 4 is insufficient. All

the votes are in. 20 percent voted high, 80 percent voted moderate, zero voted low and zero voted insufficient. So for reliability for measure 0655, the measure passes.

CO-CHAIR YAREMCHUK: Okay, so feasibility--okay. Validity in terms of this is, you know, discussion of the ability to find that it's important. There is exclusions for medical reasons that we've talked about previously regarding this. The data measures diagnosis and treatment, so validity for this, I would say is high.

CO-CHAIR MERENSTEIN: Do you want add anything? The question is, is this the same as the other two? Do we have to--we can use the same two votes as we did for validity and feasibility, or it's different? You have to speak in.

MEMBER RAMBASEK: I would tend to feel that way.

CO-CHAIR MERENSTEIN: Any other comments or questions before we vote on validity?

We talked a decent amount about this--paper issues.

DR. WINKLER: I guess the question I would ask is do you have concerns about the fact that many antihistamines are over the counter medications in terms of capturing the data?

CO-CHAIR YAREMCHUK: And I guess my response to that is that if you're looking at the chart, the person who's going to abstract it and look at it and say I recommend antihistamine or Claritin or Sudafed or something like that, and similarly if we do it electronically, if there's a CPT code and that's your recommendation, you'll capture it that way as well.

CO-CHAIR MERENSTEIN: I think we can open voting.

MS. ROBINSON-ECTOR: Voting is now open for validity for measure 0655, and for those on the phone, option 1 is high, 2 is moderate, 3 is low and 4 is insufficient. All the votes are in. 7 percent voted high, 93 percent voted moderate, zero voted low and zero

voted insufficient. So for validity, measure 0655 passes.

CO-CHAIR MERENSTEIN: So Reva is okay with us using feasibility and usability from the previous two votes, unless someone, Kathy or Todd or anyone else has an objection to that.

CO-CHAIR YAREMCHUK: Are there any other issues that are different than what you've already discussed for these measures compared to the others? That would be the issue.

MS. LUONG: So on the record, we are applying the same feasibility and usability voting from the previous measure onto this one, and now we will be voting on overall suitability for endorsement, for those on the phone.

MEMBER CARNAHAN: Didn't the antibiotic one have a high usability, and the steroid one have a low usability?

CO-CHAIR MERENSTEIN: We'll put it to vote.

MS. ROBINSON-ECTOR: Voting for overall suitability for endorsement for measure

0655 is now open, and for those on the call, option 1 is yes, and option 2 is no. All the votes are in and the voting is closed. 100 percent voted yes, and zero percent voted no, so for recommendation for overall suitability for endorsement, measure 0655 passes.

CO-CHAIR MERENSTEIN: So we're going to switch now to get some new developers. Thank you. Just so you know, it's appropriate testing for pharyngitis. Start whenever you're ready.

MR. HAMLIN: Good afternoon, I'm Ben Hamlin, I'm the Director of Performance Measurement, NCQA. Is this better? Sorry. I'm Ben Hamlin, I'm the Director of Performance Measurement at NCQA.

DR. BARTON: And I'm Mary Barton, Vice-President for Performance Measurement.

MR. HAMLIN: So this measure is a health plan measure of percentage of children two to 18 years of age who are diagnosed with pharyngitis and dispensed an antibiotic, who then received a streptococcal test. So basically,

effectively looking for appropriate testing following a diagnosis and administration of the antibiotic for the measure.

The measure's been in HEDIS for several years. It was developed some time ago, and I believe there's a number of--there's some more recent performance data in your forms I believe, I don't know how much more you want us to go into now.

CO-CHAIR MERENSTEIN: Okay, this is mine, so for the evidence, both the IDSA and the American Heart Association guidelines recommend that a throat culture or a rapid test is done before treating. So the evidence to me is the biggest issue, the reason is because the American Academy of Family Physicians, the ACP and the CDC recommend, unless developers know differently, on a decision rule, where you get five points and if you have zero to one point, it's a virus, you don't do anything.

If it's five, you treat with an antibiotic without testing, and then if it's two,

three and four, you would follow these guidelines to do testing. So that is the only -- the major problem I have with the evidence, or the question I have with the guideline considering that there's other groups that don't recommend it. I don't know if Jackie has anything to add. Any other questions or discussions about the evidence behind actually doing this? Developers, you have anything with that?

MR. HAMLIN: I should have added one more thing, that it is an administrative claims-based measure, so many of the decisions that you were discussing are not able to be captured in claims.

MEMBER FRIEDMAN: I'm sorry, so the evidence shows that you should treat them this way or you shouldn't treat them this way?

CO-CHAIR MERENSTEIN: Well, that's the debate. The IDSA and the American Heart Association recommend not treating with antibiotics unless you have a positive quick strep or a positive culture. The other recommend

use a decision rule, where you get different points for different clinical characteristics, and if you have zero to one, then you say it's a virus; you don't do anything.

If you get two, three or four, then you would follow this guideline and do the testing, and if you get five you just automatically treat without testing. And so I do think you get that from the administrative data. You would get if they were treated with antibiotic and diagnosed with pharyngitis without testing. You would get that from this data set.

MR. HAMLIN: Yes, you would get the diagnosis on the antibiotic, you would not be able to attract the decision, the scoring you mentioned through the --

CO-CHAIR MERENSTEIN: Exactly, yes, but you would -- that's my point. You would get -- they would be considered doing it incorrectly. If the physician was following the CDC Academy of Family Physicians and the patient got five points, they would treat them with antibiotic

without testing, and then guideline would say they're doing it incorrectly, right? That's correct.

MEMBER MADONNA: Just for clarification, so the order here is present with pharyngitis, strep test, antibiotic if the strep test is positive, correct?

CO-CHAIR MERENSTEIN: Yes, or a culture which--

MEMBER MADONNA: Yes, either way. Then why is it written the way it's written? Why doesn't it say the percentage of children blah blah blah who are diagnosed with pharyngitis receive the group A strep test for the episode, and then were dispensed an antibiotic if that test were positive. Isn't that what we're getting at?

MR. HAMLIN: So the order is a little different, the order is actually pharyngitis, antibiotic, strep test in the claims, because we do not -- we're not looking for the result of the test itself, we're just looking at if the

antibiotic --

MEMBER MADONNA: Okay, so it's --

MR. HAMLIN: -- was dispensed, the test was performed.

MEMBER MADONNA: Okay. That's actually an important point. They don't get that -- we'll talk about that later too. That's not an evidence issue, but yes.

MEMBER RAMBASEK: Do we have any evidence of what percentage of children with pharyngitis are going to have five points? Because that might be an unusual presentation and if it's only three percent, it might not affect it as much as we would worry.

MR. HAMLIN: We're looking for it right now.

MEMBER FRIEDMAN: So even if it's a small number, I still don't want to be dinged for doing what I think is best clinical practice. If it's one percent of the time, you're going to ding me for doing what I think is best practice for this patient? I'm just --

CO-CHAIR MERENSTEIN: It's not just what you think, what's been recommended by a major group, so.

MEMBER FRIEDMAN: Can you elaborate on that, whether you agree with what Dan's saying or disagree or --

DR. BARTON: What is a little bit ironic is that the purpose of the AFP and CDC decision rule is to make sure that antibiotics are not being used in children who should not be treated with antibiotics, right? That's the reason for defining the decision rule with five gradations, is to say reserve for the people with the highest prior probability of strep pharyngitis, reserve those people to get the antibiotics.

Our measure is coming at it from sort of a perpendicular direction, but also has the same goal, which is to make sure that children are not being treated with antibiotics unnecessarily. I think from the point of view of a health plan measure, it's often the case that a

claims based measure is an approximation of what a clinical guideline in a physician's office might say.

And that has been something that has worked very well over the last 25 years to hold health plans accountable for providing good quality care to their members, and it's not always something that has been, as you say--maybe it was not you, Dr. Friedman, someone else who used the word dinged, you know, so it's not necessarily the intention of NCQA to say that the health plan should be assessing each particular case necessarily, as though there was a clear line of right and wrong.

This is, again, there are many, many children who present with fever and sore throat. And so that's what this measure pulls on, are those claims in order to say did you mostly do the right thing at the right time, which is test for, you know, not four out of five necessarily, I don't know if the distribution of kids is in the decision rule, but the majority of kids that

you did the right thing and tested before presumptively treating.

MEMBER FRIEDMAN: Could you get around that by doing what they did for the last few measures and say avoidance of inappropriate use, and make it an overuse measure? So the percentage of children who are diagnosed with pharyngitis who didn't need antibiotics and were not given antibiotics. So that's basically what you're trying to do, is avoid overuse of antibiotics, is that correct?

DR. BARTON: Yes, but I'm not exactly sure how that -- how this measure -- how inverting or emphasizing one different part of the numerator or denominator would solve the problem that you're talking about, so I might need a little more coaching on that.

MEMBER SCHACHAT: I just interpreted this as not necessarily a measure trying to avoid antibiotic overuse, but a measure trying to encourage strep testing, which presumably would reduce inappropriate use, but I'm okay with the

whole thing. I think the concern about being dinged relates to how this gets used, because this gets used different ways by different people. So like Medicare, they'll just pay you more or less if you report, and it doesn't matter what your grade is, as long as you're reporting. So you don't get dinged there.

Some other group may be looking at your quality performance, and you might effectively be being dinged, but then if the number of -- five out of five kids is similar across all people, everyone's getting dinged equally. Probably though, different providers have different case mix; that is, maybe doctors in the ER have seen more five out of fives and they get dinged more than family practice people in a simple office; I don't know. So it would be nice if it could account for that.

CO-CHAIR MERENSTEIN: Any other questions or comments? I'm sorry, I don't have much more information to--

MEMBER YOUDE: I have a question on

the denominator statement. So looking at it, it looks like we're including children that carry strep, anytime you test those kids, it's going to be positive. So why not exclude them?

MR. HAMLIN: Again, we're not looking for the actual performance, the actual result of the test, we're just looking that the testing was performed, and there is probably a case where, you know, to be made for not testing a certain subset of children, but I think it probably increases more noise in the measure at the health claims level than it probably does solve the problem of children carrying strep.

MEMBER YOUDE: I guess when I look at it, if I know my patient has strep, and trust me, I'm speaking as a non-ENT here, so if I'm using different words, I apologize. But if I know my patient has strep, I have no reason to test them. So wouldn't I be increasing the cost of healthcare by testing kids I know don't need testing?

MEMBER LYNCH: How often does that

happen, though?

CO-CHAIR YAREMCHUK: And I'm going to say that's probably the question. I mean the incidence of carriers I think is very low, would you say single digits? So very, very low, and I'm just looking at what the mean is in the Medicaid population, it's like 66 percent of people doing it, and for the commercial it's 79 percent, so you gap is, depending on--20 to 30 percent, and it's clearly not what the percent of carriers are.

MR. HAMLIN: And I also want to add to that that the end there is health plans, not patients when you're talking, so there's 400 plans, not 400 patients in those.

CO-CHAIR MERENSTEIN: So I guess we should vote on the evidence and going forward with this.

MS. ROBINSON-ECTOR: Voting for evidence for measure 0002 is now open, and for those on the phone, option 1 is high, 2 is moderate, 3 is low, and 4 is insufficient

evidence. All the votes are in. 13 percent voted high, 53 percent voted moderate, 27 percent voted low, and 7 percent voted insufficient evidence. So for -- oh yes, so the measure passes.

CO-CHAIR MERENSTEIN: We're going to go on for opportunity for improvement. Kathy already took what I was going to say. One, so there is differences between the Medicaid and the commercial rates. The rates still aren't very high. They report that the level of performance indicates the clinical guidelines addressing testing for our group A strep infection prior to treating pharyngitis with antibiotic are not followed in about 25 percent of cases on average, and this is likely an underestimate.

So there's no question that there's overuse of antibiotics and inappropriate usage, so I think that was a huge indicator. If we think this is a good measure to measure it is the question, but there's no question there's room for improvement I think. Jackie, anything to

add?

MEMBER RAMBASEK: Just clarification. You said 25 percent of patients who present with pharyngitis get testing?

CO-CHAIR MERENSTEIN: No, no. 25 percent are treated with antibiotic -- are not -- wait, let me -- prior to treating pharyngitis with antibiotic are not followed, so 25 percent of the time, they're just treating them without testing. Is that--yes, so theoretically, they have a clinical guideline you have to score a five, or it could just be they're just treating people with viral pharyngitis.

MEMBER RAMBASEK: But we didn't clarify what that number is yet, did we?

CO-CHAIR MERENSTEIN: No, we don't know what that five is. It's definitely not 25 percent though, you would think.

MEMBER RAMBASEK: Probably five?

CO-CHAIR MERENSTEIN: Yes, that would be my guess clinically, but that's just an anecdotal clinical guess. Any questions about

the room for improvement? We can vote.

MS. ROBINSON-ECTOR: Voting for performance gap for measure 0002 is now open, and for those on the line, option 1 is high, 2 is moderate, 3 is low and 4 is insufficient. All the votes are in. 40 percent voted high, 47 percent voted moderate, 13 percent voted low, and zero voted insufficient, so performance gap for measure 0002, measure passes.

CO-CHAIR MERENSTEIN: So going on to reliability. This is health plan and administrative data claims, Jackie is going to talk about some of the exclusion criteria and the questions about those. The one thing it was clear is the data were very consistent over time, and it didn't really change for the last 10-year data, so it looks like the rates, they're getting what they're recording. You have anything to add about that?

MEMBER YOUDE: I just had a few questions about the exclusions. One, and we may have touched on it earlier, so what about kids

who get the rapid strep test, but we don't wait for the results, we just--we give them antibiotics regardless. How does that work into it, and then I've already brought up the one about those who carry strep.

MR. HAMLIN: So in the first scenario, the patient who received the antibiotic and was tested will meet the measure criteria. I should also mention there is a separate measure that's looking at using antibiotics in children with upper respiratory infections. That is a separate HEDIS measure, and that is getting more at the overuse of the antibiotic issue. So -- and again, we know of the carrier issue, but it's--

MEMBER YOUDE: Okay, so I guess this goes back to our original discussion that this is truly just about testing, and it's not about the result of the test, so if this kid had a virus, and we tested him anyway because hey, we were on like that two or three number scale that we had just talked about, turns out kid has a virus, but we gave him antibiotics just in case, this person

would still be okay because they did the test?

MR. HAMLIN: We hope it would lead towards the next time do better rule, and--

DR. BARTON: Or you call the family and tell them the test was negative and stop taking the meds.

CO-CHAIR MERENSTEIN: Yes, but that's the major problem, that we just don't know that. This test doesn't tell us that. It doesn't tell us.

MEMBER STEIN: Can you comment on, you know, one doc -- it's possible one doc codes for the pharyngitis, and another doc codes for the -- prescribes the antibiotics, the data sources you're using, is that linking the same doc, or could someone code for one of these and then get dinged because a different provider is doing the other part?

MR. HAMLIN: So the claims that we are using are linked to the episode, so whenever the patient appears on the claims record for that one patient, regardless of which provider actually

submitted the claim around that date, those would be linked together by the health plan itself, so--

MEMBER STEIN: But in real practice, they may not be linked, right?

MR. HAMLIN: We don't get this at the practice level, the plans are responsible for summarizing the claims around each episode to define the episode of care, and they submit the final aggregate to us, so we don't have that level of measurement at the actual practice level of their behaviors. It's just -- it's a plan level measure, not the practice level measure.

CO-CHAIR MERENSTEIN: Any other questions about this, about the reliability? And you had some questions about the codes, right?

MEMBER YOUDE: I did have a question about the codes. So I am not an ENT, obviously, and so when I got this lovely measure about pharyngitis, I did the first thing I knew what to do, and that was go to my laryngologist friends, and I had them take a look at the codes because

the codes are not something that I'm familiar with, and they had brought up that at least three codes that they looked at on first glance were missing.

Had to do with acute pharyngitis and another one, and there's another one, but it was at least three known codes that are commonly used were missing. How did we come up with the codes, where do they come from, and do you think that there are potentially any missing, or is this an exhaustive list?

MR. HAMLIN: So the codes are initially developed when we test the measures, we look at the series of diagnoses and then we have several modes of updating those code lists every year, and we have a policy clarification support system where people provide us -- suggest missing codes.

Through implementation of the health plans, they can suggest missing diagnoses that we might want to code, and then there's a regular -- there's several expert reviews with a coding

panel and a clinical panel that review the codes that need for each of these. And so I'm not familiar with the codes you mentioned and why they're not necessarily included, but it's usually sort of -- it's part of the process, they either were not found to be reliable enough across the different health plans, or they may be missing for some other reason because of the way the coding practices are.

CO-CHAIR MERENSTEIN: Any other questions or comments before we vote on this? Open voting.

MS. ROBINSON-ECTOR: Voting for reliability for measure 0002 is now open, and for those on the line, option 1 is high, 2 is moderate, 3 is low and 4 is insufficient. All the votes are now in. 7 percent voted high, 73 percent voted moderate, 13 percent voted low, and 7 percent voted insufficient. For reliability, the measure passes.

CO-CHAIR MERENSTEIN: So we move on to validity. And so I had some previous questions,

I have some more, but I think for validity, it's definitely getting at what they want to get out of questions if that's a big thing, but what they did is they look at five different plans, geographically varied, and did medical record extraction, and they found that there was agreement 86 percent of the time, so I think that's a pretty high agreement.

And if Jackie has anything else to add about that? So I think that with the -- they're answering the question they want answered, we've got to decide if that's the feasible question, but -- and for validity, I think they are. Any questions about that? All right, we're going to vote.

MS. ROBINSON-ECTOR: Voting for validity is now open for measure 0002, and for those on the line, option 1 is high, 2 is moderate, 3 is low, and 4 is insufficient. So it looks like we're missing one vote from someone in the room, so if you all could re-vote, please. Still waiting on one person. Thank you. 27

percent voted high, 73 percent voted moderate, zero voted low and zero voted insufficient. So for validity for measure 0002, the measure passes.

CO-CHAIR MERENSTEIN: So maybe I misinterpreted, I just got whispered, but I was going to talk about the feasibility, and you're saying how easy it is at the clinic. So it's easy to do an admin thing, I guess since maybe I missed it I'll add it now. My main worry is--the whole point of this testing is to make sure people use antibiotics appropriately.

And I guess I'm not convinced that the way you're doing it is -- that's answering the question, all you're really getting is for people to make sure they order the quick strep or the strep culture, and then they can do whatever they want, and they're considered appropriate care. And to me it seems like today, just like you were talking about at Kaiser and other EHRs, there's a better way to get the data we want to get appropriate use of antibiotics. But I think I

brought that up at the wrong place.

CO-CHAIR YAREMCHUK: I agree, and I think we've touched on that throughout the discussions, but I agree entirely.

DR. WINKLER: I think that's the appropriate--usability and use is probably the appropriate place to talk about that. So feasibility is really about the data source, how the data is generated and transmitted and implemented.

CO-CHAIR MERENSTEIN: So that, I mean I think it's -- there's no question it's generated by the plans, it's administrative data, it's been very consistent over time, there's really no question about that I think. So I guess we can vote on that. Oh there is a question, I'm sorry. Josh?

MEMBER STEIN: Is it showing any improvement over time?

CO-CHAIR MERENSTEIN: It is a slight --let me pull it up.

MEMBER YOUDE: If we're looking at the

commercial rates and Medicaid rates, is that what you're pulling up?

CO-CHAIR MERENSTEIN: Yes.

MEMBER YOUDE: If we're looking at the mean, it stayed within about -- it stayed within one percent for the commercial rate. If we're looking at Medicaid, we've stayed within two percent for the mean since 2012 to 2014 in both types of insurance. So this may relate back to our earlier conversation about, you know, we have a process metric, we keep it in place because it holds the gains, but my question about this then comes to what's the expectation?

CO-CHAIR MERENSTEIN: I guess we have to vote on feasibility, and that's a usability question I guess. So let's just vote on this, and then we'll talk briefly.

MS. ROBINSON-ECTOR: Voting for feasibility for measure 0002 is now open, and for those on the line, option 1 is high, 2 is moderate, 3 is low and 4 is insufficient. And we're still waiting for those on the phone to

vote.

MEMBER LYNCH: I voted.

MS. ROBINSON-ECTOR: Thank you. Great, so all the votes are in. 33 percent voted high, 67 percent voted moderate, zero voted low and zero voted insufficient. So for feasibility for measure 0002, the measure passes.

CO-CHAIR MERENSTEIN: So usability, you heard my issues, and Jackie I don't know if you want to repeat anything, but you heard Jackie's issues too, with the problem with that.

DR. WINKLER: One question I would ask the developers, since they are the implementers of this measure, a question that's coming up on these measures that have been around for a while and have been used, particularly since we have data over time, is how are you assessing the impact of the measure? What do we know is really the benefit of this measure being in place in health plans? How--are you making any assessments of that impact in terms of what's happening in the quality sphere? Is it making a

difference, if so, how?

DR. BARTON: Our committee on performance measurement has shown an interest in having such information, and we agree. We have been trying to understand how, of course, the span of 70 plus measures that we use in HEDIS so that we cannot necessarily commit a large amount of resources to each one each year, but we are trying, in a rotating fashion, to gather tales from the field to help both our committees that volunteer their time to help us keep these measures up to date, in the loop. I think that there's no question that measures that have, for example, found their way into the Stars program have had a rapid acceleration of their improvement, and this is not one of those measures.

However, there are currently efforts afoot, including collaboration between NCQA and NQF to work on the issue of antibiotic overuse, and so I think that as we think about health plans accountability, which is only one part of

course of the entire issue of antibiotic overuse, that this is a valuable tool for health plans to use.

And we also have in the upcoming measures that will be used in the quality ratings system for the marketplace plans, which of course won't be public for another year or two, this measure is included in that group. It will be interesting to see how it plays out in that setting and additionally in Medicaid over the course of the next few years as the Medicaid and the marketplace settles out.

CO-CHAIR YAREMCHUK: I guess my only question is this really doesn't measure antibiotics, because you don't know if it was prescribed or not prescribed. You're assuming that if it comes back negative, they won't get a prescription, but you really don't know that. So the question of over-use and under-use is hypothetical, I'm going to say. And we were just talking that in terms of the change as far as measurement, it hasn't budged a whole lot in a

couple years.

DR. BARTON: I'm not sure, but I think measure 002 means that it was first endorsed by NQF a very long time ago.

MR. HAMLIN: Again, at that paired measure, it actually does look at antibiotic overuse as a very similar slow rate of improvement. So I think it probably is something to do with the area of antibiotic use.

MEMBER YOUDE: Is there--like when I look at a process measure, and I'm like okay, what do I need to do, when do I need to do it, that type of thing, what am I being held accountable for? When I look at this, to me it says you need to test. You need to do more tests and test, test, test, test, test.

Is there any concern that we could be increasing health care costs because we're doing more and more and more testing? Especially given the evidence seems to conflict with practice patterns where on that decision point scale, if somebody's telling me to test and I'm going to be

reimbursed on tests, I'm going to test. But it might not be appropriate testing.

MR. HAMLIN: The cost of this particular test pales in comparison to the cost of the antibiotics that are being dispensed, so I think --

CO-CHAIR MERENSTEIN: I didn't hear you.

MR. HAMLIN: I said, the cost of the test itself is actually -- it pales in comparison to the cost of the antibiotics that are even being dispensed, so I think that we need to think of it as low cost test that is not going to be driving over-utilization. And again, it in the context of the thinking about prescribing antibiotics in the first place for these kinds of ---

CO-CHAIR MERENSTEIN: Although I would argue that because it's so inexpensive that now you just told every clinic just test before the doctor walks in the room, and then you've got your answer because it's not going to cost you a

lot of money anyway, just test and then the doctor, whenever the doctor does, the doctor's fine.

MR. HAMLIN: But if you've already prescribed the antibiotics, then you can to think about the long-term consequences of antibiotic overuse. So it's the chicken and egg there in that argument.

CO-CHAIR MERENSTEIN: Well yes, I'm not arguing you should give the antibiotics, I'm saying that you still would look on the measurement if you tested the quick strep before I even walked in the room.

MR. HAMLIN: Well, the measure only takes in if you've been dispensing antibiotics in the first place. So you would have already had to administer the antibiotic and then give for the test, and then you'd be in compliance for this particular measure, so.

DR. BARTON: You're right.

MR. HAMLIN: You don't need to convince a doctor to stop antibiotics, you just

need to test, that's the point.

DR. BARTON: No, no, you're right that the scenario you described would result in perfect performance on the measure, but it would be an extremely inefficient way to do that. Because you'd end up testing a whole lot of people you would never even consider giving antibiotics to.

CO-CHAIR MERENSTEIN: Any other questions or comments about this? I guess we'll vote on usability.

MS. ROBINSON-ECTOR: Voting for usability and use for measure 0002 is now open and for those on the call, option 1 is high, 2 is moderate, 3 is low, and 4 is insufficient information. All the votes are in. Zero voted high, 53 percent voted moderate, 40 percent voted low, and 7 percent voted insufficient. So, this would fall in the gray zone.

DR. WINKLER: This is a consensus not reached decision, but this again is not a binding one in terms of moving forward. So I guess the

question is based on this, does this prompt any further discussion from this committee? Certainly, we will anticipate feedback from public comment, but also you'll factor in your evaluation of the criteria as you go to your next vote on suitability for endorsement.

CO-CHAIR MERENSTEIN: I guess maybe I've already said this, I think it is a really important measure if they can do it correctly, and I just don't think it is done correctly. So that would be my vote for the next one too. So I don't want the take-home message to be it's a bad think to check. I think it is a really important thing to check, but I'm just not sure it's being done correctly.

MEMBER FRIEDMAN: Can you make some recommendations on how you would make -- and I already addressed this myself, that I don't think -- I think it could be stated differently as well. So how would you state it to where it would be done better?

CO-CHAIR MERENSTEIN: I think you need

a way, a measurement that says when the test is positive, is an antibiotic given and when the test is negative, is an antibiotic not given? Somehow you need to get that data. Because there's not that many people that argue a positive test doesn't need an antibiotic.

MEMBER RAMBASEK: Are people going to come back and say they can't collect that data? Or is that going to be, if they write that way, is that going to be the argument pushback, it's not usable because we're not going to be able to get the data on the test result?

MR. HAMLIN: The test results are not available, so that's --

CO-CHAIR MERENSTEIN: You can't get that from laboratories?

MR. HAMLIN: Not the direct laboratory feed, no. We get the paid claim for the testing form and the result itself, the numeric result, is not available in that data.

CO-CHAIR MERENSTEIN: Even in a place like Kaiser Data?

DR. BARTON: If we were only examining Kaiser, we would be in a very different situation.

CO-CHAIR YAREMCHUK: I have a question about that. Because for diabetes, for example, hemoglobin A1c, you actually have to get a level, LBL, you actually have to get a level. So, there's got to be some way that you get the levels to know if the patient is in or out on those things. And so for this, this would be a positive or negative. So I guess I'm wondering when we talk about that, you're talking about paid claims versus results but for these other measures, you get results.

DR. BARTON: That's an excellent point. We do require a certain number of measures that require chart review. So, anything that requires a test result or an evaluation result, like blood pressure result, requires actual chart review. And health plans then sample from among their eligible patients to report a rate. The advantages and disadvantages

of an administrative claims measure, health plans are extraordinarily loathe to add another chart review measure to what they say is a considerable burden of chart review that they do each year for HEDIS. When we have an administrative claims measure, we feel confident that it can examine performance, not just in a sample of patients, but in all the eligible patients. So there are pros and cons to each approach, but we are absolutely limited in the number of chart review measures that we can require plans to voluntarily participate in, and this has not been the public comment that we've had. The expert panels that we've convened have not indicated that this measure is one that would be so improved by chart review that they would think it worth doing that. At least that has historically been their point of view.

MEMBER RAMBASEK: So basically we said it's not very usable and then we suggested a change and you're saying the health plans would say our change is not very usable?

MEMBER STEIN: Can you explain a little more? If it's a health plan, shouldn't they have access to the lab results?

MR. HAMLIN: Sometimes yes and sometimes no. Reliably across all the different health plans that report, no. They do not always have access to the labs. Different designs obviously have much more access to that indiscrete data than others, but usually when they go and do the medical record review, they have to the paper medical record and look up the 411 charts to find those results. And diabetes is 528 so, it's a considerable number of records that they have to actually go in and look in the actual record itself for the lab result, not from a laboratory feed or from some kind of electronic system.

CO-CHAIR MERENSTEIN: Any other comments? I guess we'll do the final vote.

MS. ROBINSON-ECTOR: So the vote for overall suitability for endorsement for measure 0002 is now open and for those on the line,

option 1 is yes, and option 2 is no. All the votes are in. 33 percent voted yes and 67 percent voted no. So for recommendation for overall suitability for endorsement, measure 0002 does not pass.

DR. WINKLER: I'll just say that this is a measure that's been around a long time. It's being used out there. I can anticipate a significant number of comments during the public comment period, and so we will revisit this issue more than likely, based on that.

CO-CHAIR MERENSTEIN: Do you have a comment?

MEMBER SCHACHAT: Is there a process for revisits? I mean if you get a bunch of people complaining, is there going to be another meeting and another vote?

DR. WINKLER: Not necessarily another meeting. That's the point of public comment is feedback. So, you will reconvene by conference call in August to review the comments and I'm just predicting that this is likely to prompt

some comments and that would then be on your agenda for revisit.

CO-CHAIR MERENSTEIN: Todd?

MEMBER RAMBASEK: So, I guess the basic push-back that is being given is that people will order the test and it will be negative, and they'll prescribe the antibiotics anyway is what people are worried about? Do we have an idea of how often that actually happens, that people are using the test as just, you know, and not respecting it, not listening to it?

CO-CHAIR MERENSTEIN: Happens every day in my clinic. I don't know about other clinics but it happens every single day in my clinic. It's ordered before you get there. So I would say that sometimes it's ordered when I think it's a virus and should never have been ordered, and then other people are going to do antibiotics no matter what they order, so they don't care what the tests are. They say, you know, looks positive to me, I'm going to treat it. So, I imagine it's going on.

MEMBER RAMBASEK: So, I have anxiety about that. Is it being ordered by like a resident or fellow or by a non-M.D.?

CO-CHAIR MERENSTEIN: By the administration. The nurse orders it before you even get there. The nurse, you know--

MEMBER RAMBASEK: Okay, so, the nurses are allowed to order tests?

CO-CHAIR MERENSTEIN: Yes, the nurse is already doing the test before you get there, yes. Considering all the tests they order, urine, x-rays, things like that, I'd say this is number one that I would say is a mistake.

MEMBER RAMBASEK: Are there issues about having the test paid for if it's ordered and you found it wasn't needed, you're supposed to cancel the charge? How does that work?

CO-CHAIR MERENSTEIN: Maybe someone else can answer that, I don't know. I mean, I think they don't order if someone doesn't say they have a sore throat or a runny nose or some kind of respiratory---

MEMBER RAMBASEK: I just work in a different sphere. All my patients have Medicare and I don't have nurses. And so, if someone orders a test that I decided I didn't need, it's a big problem, and all kinds of bells go off and we have to discuss, how do you even order that, you shouldn't have pressed the button and it's complicated for us to cancel the charge when we cancel the charge. But when the charge is cancelled, that raises red flags about cancelled charges because--anyway.

CO-CHAIR MERENSTEIN: So unless there are any other comments, we'll go on to some ophthalmology issues.

MEMBER FRIEDMAN: I just have one more quick question for Reva. So, with the public commenting period, how often does that change the voting?

DR. WINKLER: I wouldn't say often, but we will ask you to very seriously consider the comments that you get. But it can and has changed.

MS. LUONG: And to add on to that, if you have any specific requests for the measure developer, either now or for the June 22 call or for the post-comment call in August, please let us know and we can work with the developer to come up with a response, if possible.

DR. WINKLER: Okay, and if you didn't think we had a lot of information for you, I'm going to add a little more. In the eye care measures, among the group that we're going to be reviewing this afternoon and tomorrow morning, we will now be seeing a new animal called a true eMeasure, and I want to be sure that everybody's clear what we mean by that. eMeasures are a type of measure that are very specifically defined, that use electronic health records. But the eMeasure is specified in a very specific fashion. It is not just using--querying your EHR and getting the data. That is using your medical record as you would a paper medical record, if it happens to be electronic, that's nice.

An eMeasure is very specifically

written and specified in a format that's an industry standard. It's called HQMF. It also uses very specific definitions around the quality data model, and it also has the value sets. In other words, the acceptable values for each data element through the National Library of Medicine's Value Set Authority Center. So, we aren't talking just measures that can be done in your EHR, we're talking about measures that are specified in a very specific fashion. And so those HQMF specifications have been included with the submission for those measures that have an EHR version as well as say perhaps a registry and maybe even a claims version. So, we're seeing sort of a mixture of both.

MEMBER SCHACHAT: So, the EHR vendor would have to opt to insert that electronic structure into the electronic medical record for it to work?

DR. WINKLER: Perhaps. There needs to be some ability of the EHR to utilize the appropriate HQMF specifications. So go to the

next one. I just want to show you, if you haven't seen them, these are what they look like. This is the output for the measure authoring tool which is a sort of facilitated way of creating these HQMF specs. This is the human readable file. If you open a couple of the other files, you'll find it's totally computer language and I hope you enjoyed it. So move on to the next one. So if you go through the human readable side, you see all of the different elements of the data elements--and go on to the next--and there were some--next one. Here again are the QDM data elements and the last one are the value sets. So, I just wanted you to be aware that when we talk about an eMeasure, it means it has this particular set of specifications. All right, next slide.

Development of eMeasures has been an evolutionary journey over the last, I don't know, six, seven, eight years, whatever it's been. Really got launched when the HITECH Act was enacted in 2009 to promote the adoption and

meaningful use of HIT. I think everybody's probably familiar with the meaningful use concept. And so, first IONC and now at CMS there have been the incentive programs to develop and adopt, to adopt and utilize, in meaningful fashion, programs from the federal government and most commonly called meaningful use. They have them for hospitals, it is for eligible professionals as well.

The purpose is to gain experience with the development use of eMeasures and use of measurement within an electronic health record, using these particular specifications, the HQMF. So as part of getting going with that, there were eMeasure specifications created for existing quality measures; so they're called retooled measures, all right? So there was a regular measure, call it claims, paper, registry, whatever, but then the eMeasure specifications were created. So when we refer to retooled measures, that's what we're talking about. And so this whole program is meant to encourage the

development of more and more eMeasures hopefully that will begin to use the unique capabilities of EHRs.

But we are still in the midst of this journey and transition, all right? So, I think the hopes and dreams of pretty much everybody in the EHR space is that we'll get someday to a point where they really do a lot of wonderful things. And we're not totally there yet. The development and testing of these eMeasures has been quite challenging. The testing of the eMeasures is hindered by a lot of things, but one is the limited use in the field. For instance, PQRS allows measure to be submitted through EHRs but in the most recent data in 2013, only 66 eye care specialists, probably ophthalmologists, submitted data to PQRS using EHRs. The vast majority of them used it through either the claims or registry option.

So we don't have a lot of big data working. There are some folks who are particularly working the registry space,

developing it, but we are still in transition. And in fact, in the meaningful use program, a lot of the problems with data is the fact the eMeasures are reported by attestation. They're saying yes, I can do it. But the actual data is not transmitted, so there's no collection of a data set that someone can do some statistical analysis for liability and validity so, we're really stuck. So, these are the challenges that my colleagues in the measure development world are dealing with. So we're going to find sort of an awkwardness and a difficulty in trying to understand how well these measures, the eMeasure versions meet the criteria. Next slide.

Okay, so NQF's current criteria for the evaluation of eMeasures is we do look at these measures distinctly. So even though you had one measure worksheet that had information for both the registry measure say, and the eMeasure, we will be considering them differently because the issues around reliability and validity, the issues around feasibility, and use

and usability are distinct for the eMeasures versus, say, a registry measure. So, we will be doing those independently.

We have provided--our staff has provided a technical review of the eMeasure that's included in your worksheet so, that will apply to the eMeasure. And in general, the eMeasures are expected to meet the same criteria. So it's the same things you've been going through. There are some special applications for eMeasures, in a perfect world, we really would like to see testing for liability and validity in more than one EHR system, or a system from different vendors, so we can see how well it works. That's very hard to achieve, actually. And then also the feasibility assessment that addresses the data elements and the measure of logic. And so that's where there are some specific issues for feasibility for these measures.

As it turns out, we've got six retooled measures that are on the table for

evaluation alongside their endorsed, sort of original version, either registry or claims. And so we will evaluate them distinctly. We acknowledge the transition, we acknowledge the challenges, and so we've made some sort of adaptions, if you will, to the criteria for these particular retooled measures that have this long history of trying to evolve with the rest of the industry. So, next slide, please.

So, these retooled measures can actually come into the same number because they are used in federal programs under the same number, so that's why they're all kind of grouped together, but we will look at them separately. I know that it's clumsy, but at this point we really do want to talk about the measures independently even though we do carry them under the same number. We are looking for ways of being able to indicate the two different versions in the number without totally upsetting the federal rule making process that actually indicates the number of the measure when it's

included in federal programs. So we will be looking at them independently.

Absolutely, totally new, since so many of the developers are struggling to get these measures tested in more than one EHR system, new guidance that is allowing an option for testing of these measures, the retooled measures, in a simulated data set. There is a tool that CMS has created, MITR actually created it, that allows you to create a simulated data set and run the data that you have from your EHR to test the measured logic and be sure that it operationally can create and give you an answer. And then there will be the ability to compare the results from the automated EHR versus whatever you programmed for your simulated data set. So we've just been advised the major developers of that new guidance, and so they haven't had the opportunity to do it. But in discussing it with them, that is certainly an option they may choose to take and they might be able to get it done like while we're out for comment.

And so this is something that you can consider as on the table and under the conditions that would be acceptable to recommend the eMeasures for endorsement. Is that my last slide? I can't remember. Okay. So with that sort of introduction about eMeasures, does anybody have any questions about what I just said? So we do understand that there are these-- that eMeasures are very specific; it doesn't mean just doing your measurement in your EHR. Okay. Alrighty, with that, eye care measures.

DR. RICH: All right, I think that I'll start with some introductions and then I'll give you some general thoughts about the role of the Academy and what we are doing here. So, my name is Bill Rich, I'm Medical Director of Health Policy for the American Academy of Ophthalmology. I practice in Northern Virginia, and I'm also the Chair of the Measure and Development group for our IRIS clinical registry. Flora?

DR. LUM: I am Policy Director of Quality Care and Knowledge Based Development and

I was actually the original staff on these eye care measures when they first came through with Reva. Yes, you remember that?

MS. CHAVARRIA: Thank you. I'm Elvia Chavarria, I'm with the American Medical Association Physician Consortium for Performance Improvement and I will--if necessary, I will invite other PCPI colleagues to answer any questions that you might have, that they might be better suited answer.

DR. RICH: Thank you. On behalf of the American Academy of Ophthalmology, I'd like to thank all of you for taking time away from your families and practice to be here, the staff for their diligent work, and after listening to the description of eMeasures, I'm glad there are a lot of ENT docs in the room, because people are going to start spinning the number two pencil in their ear. So what we're going to be talking about are several measures that deal with the leading causes of blindness in the Medicare population, macular degeneration, diabetic

retinopathy, and glaucoma. We're going to be talking about a cross-cutting communication measure dealing with sending a report to the primary care docs. We're going to be looking at an intermediate outcome measure in diabetic control, and finally, two patient-centered outcome measures that will be publicly reported next year.

We look at the ability to improve quality and better outcomes as a continuum from-- you have to start with great science. Number two, you have to have dissemination and education. And that's where we've been in the past. Now we're moving into an area where we can hopefully measure and then a registry further give rapid feedback to performance improvement. So if you look at some of the outcomes, the process measures, you have three before you that we'll be discussing later in diabetic retinopathy, macular degeneration, and glaucoma. So the question is, how effective has that been, these process measures, you don't look at the

retina, you actually stage disease. So, if you have cancer of the throat and tongue, you get CAT scans and a lot of other work-up, we can actually look in the eye, educate docs on the staging of the disease to incorporate better outcomes and appropriate therapy.

How effective? Well if you look at the gold standard for clinical trials, it was diabetic retinopathy study and the EDTRS study in the early 80's. The Academy then aggressively promulgated the knowledge base, we communicated with about 80 percent of primary care staffs around the country and within five years, we decreased blindness from severe macular degeneration 50 percent and we're approaching 70 percent now, and with 18 percent less resource use. So even with the education and staging, we've made great strides. We feel very strongly with measures and then testing them and then implementing and measuring in the registry, we're going to even consolidate and address the gaps in care.

The other thing I'd like to address is the perceived gaps in care and the ability to improve. All the measures we're going to be discussing today are based on 2013 claims data whereas a physician, you had to report on 50 percent of three measures, and so obviously, you're going to report on the measures you're most proficient in. And the reality is that when you look at the data, only a small percentage of physicians eligible to report actually report it. That's about 56 percent of ophthalmologists that if you throw in another 30,000 optometrists, you can see that looking at administrative claims data does not give you a picture at all as to the gap in care or the ability to improve.

Starting this year, we have to report on nine measures and when we look at the ability to report on nine measures, you'll actually see the tremendous gaps in care, and we'll be glad to address that in the registry. So I think that looking at 2015--2013 claims data doesn't give us an idea of what's happening today in 2015. There

will be no more claims reporting, you cannot report; so it's either going to be EHR reporting or through a registry. And so I think when you look at that, when we look at that data--and Flora has a lot of data--you'll be able to appreciate there's lots of gaps in care and area to improvement. So having said that, I'd like to look at also--I briefed the address at the last meeting. What happens if you actually sunset measures? And I will again distribute that article in The British Medical Journal by Joe Selby, where once you take it out, performance drops about three percent a year. Put it back in and hemoglobin A1c level performance improves 48 percent. So I think that we have to look very carefully about our assumptions about area of improvement based upon old claims data.

So the first couple of--I'll talk about the cataract measures first, and then we can move on to the glaucoma ones later on at the Chair's. The cataract measures were designed as measures of surgical competence. 20/40 was

picked because that's functional level in our society as the basis of the FDA approval for intra ocular lenses, getting driver's license and actions associated with higher levels of visual functioning. The ability to--it's not risk adjusted because we didn't feel we had the ability to risk adjust it. So, we took out everything. It's a measure of pure surgical competence which will be publicly reported. So, this is a patient-focused measure to help patients evaluate, i.e. the results of the surgery, and i.e. the surgeon's results when they choose them.

If you look at the second measure, it's the complications of cataract surgery. We turn to the OR This is something that can easily be recognizable from claims data from the EHR and this is actually something that I did wrong as a surgeon. The percentage rate is only about two percent, the complications are greatest, expensive and results of somewhere between 60,000 and 100,000 people a year. Again, that will be

publicly reported too. As far as the ability to improve, the claims data shows 163 physicians reporting the cataract measure. However, as we move toward register improvement in EHR reporting, last year we had 600,000, this year 800,000, and I think we have about 600,000 cataracts for this year alone in 2014. So you can see the distortion that will occur if you look at 163 in the administrative claims database in 2013. The same thing, and the reason why the data is so low is that those measures are registry reporting and we didn't get our registry up and running until 2014. So again, you're looking at a very small end in the 2013 database. Flora, any comments before?

DR. LUM: Yes, thanks Dr. Rich. I just wanted to reinforce to you guys that in 2013, only 138 providers, that when you see that performance rate, it only reflects 138 providers and for the complication rate, is only 77 providers. So that performance rate is high, but it probably reflects the cream of the crop who

chose to report on that measure. The second thing, I did want to address the eMeasures. As we said, IRIS registry is a certified EHR technology, which means we had to pass all the ONC tests. We have successfully integrated across 26 different EHR systems and successfully submitted the ECQM data for 2,722 physicians last year. So, I think we have shown that we have successfully created the measures, passed it to CMS's specifications and submitted those for our physicians.

The other thing I was just going to mention is, I know the ophthalmologists think a lot of these are best practice and we look at the other measures, but they are for every eye care practitioner, so we're just getting at 20 percent of eligible providers reporting these measures through claims in 2013. So as we open it up and we get more and more providers, especially optometrists as well as ophthalmologists, we will see variations in those performance rates. As Dr. Rich said, whatever performance rates are

this year, they will be publicly reported. So, as of this year, they are on the hook and they will see those rates being publically reported on the physician compare site next year.

CO-CHAIR YAREMCHUK: I guess I'm going to suggest that we move along and go with our first measure at this point in time. So, our measure that we're going to be looking at for ophthalmology, Cataracts 20/40 or Better Visual Acuity within 90 days following cataract surgery. And if we can go through the algorithm that we've done with the other ones, and Matt are going to be the lead on this?

DR. WINKLER: I'd like to interject just one thing. This is the first outcome measure we've looked at today, and I just want to point out to you that the criteria for an outcome measure for evidence is different than for a process measure. For an outcome measure essentially there just needs to be a stated relationship, sort of an actionability if you will, that there is some process or structure or

something that can be done to effect that outcome. It is not the same as required the quality, quantity and consistency of the body of evidence that you would for a process measure. So just to--it's more focused.

MEMBER CARNAHAN: Okay. So, just the thorough description of this. Patients 18 years or older with a diagnosis of uncomplicated cataract--those are all the exclusions we'll talk about later--who had cataract surgery and no significant ocular conditions impacting the visual outcome of the surgery, and had a best corrective visual acuity of 20/40 or better at distance or near achieved within 90 days following cataract surgery. And as was stated, the 20/40 number was picked based on a common DMV number, a common FDA amount used as a success target literature, as well as the Salisbury Eye Study saying if 50 percent of the functioning happened at the 20/40 level.

And I guess my only question on this was, some of the other statements made within the

measure where The Eye Care Measure Development Work Group opted to develop a measure of accountability addressing cataract surgery outcomes that relates to the safety and appropriateness of cataract surgery. I wasn't clear on the relationship there because nowadays we see younger and younger people having cataract surgery at better and better visions, below 20/40 to start. The measure is based on improving patient functional and visual outcome satisfaction and quality of life, excellent visual acuity after cataract surgery is achieved consistently through careful attention, through the accurate measurements of axial length, corneal power, appropriate selection of an IOL calculation formula. And those are all true, but I just have some concern that this measure captures that. Meaning that if I as a surgeon create a sentinel event, so I go ahead and I wanted the patient-- the patient wanted to have near vision and I gave them distance vision, and they were 20/20 distance vision best corrected,

then I succeed and I've done a sentinel event that isn't captured to get this high-quality.

And the other piece is, if I have a patient that starts out 20/30 and I have some sort of issue that comes up and they end up 20/40, I'm still successful. So I don't know if that's a wording issue within the description of the measure. I'm not even sure how you would correct that. But again, we just see more and more surgeries happening at better and better visions and so I wonder how we can make it so that this measure does what it's intended to do, which is highlight success as well as highlight opportunity for further success from others.

DR. RICH: I'll respond first. Basically, you should get excellent vision. If you don't, something happened. So that's the intent, you know, to address the issue of the target, and that's what's your talking about with the distance and near, is another measure that we've developed. I don't believe it's approved yet but it is in the registry of the QCDR where

it's actually, how close did you get to your target refraction. So, that solves the problem. We would have had a very cumbersome measure to try to address all these issues that you raised. But basically, if you didn't get to 20/40, something's wrong. And so our entire intent was not to risk adjust.

We now have a large amount of data that shows the results of about 460,000 U.S. and 370,000 in Europe, all comers with no exclusion and it's about 84 in Europe, 86 here. But that's very complicated because then you have macular degeneration, a whole bunch of things of confounding, underlying diseases that would preclude you from seeing 20/40. So the intent with this is to be a patient-focused measure that the patient can look at in public reporting and say, this person gets 86 percent, this person 96 percent. So that was the purest way that we could define a surgical competency measure. Eventually, obviously we'll be able to risk adjust this and look at people with diabetic

retinopathy and things, but we actually have no basis to do that right now.

CO-CHAIR YAREMCHUK: Okay. I guess what I was talking about and I'd asked Reva about, this is an evidence-based--the first thing that we vote on is evidence, and there's outcome and process measures. This is an outcome measure and it says report whether there are processes of care that can influence the outcome. Nothing more is required about evidence for outcome measures. So, I guess the question is, is this a process of care issue you're talking about?

MEMBER CARNAHAN: So it sounds like there are measures to do what this measure is asking to do, but this measure may not have the evidence to do what it's supposed to do. Because it doesn't actually tell you if a patient has better vision than when they started, and it doesn't tell you if the target that was hit was the one that the patient wanted. So you can do atrocious things and succeed at this measure. So is that evidence of a successful measure?

MEMBER SCHACHAT: So those scenarios you're speaking about can happen, but fortunately they are vanishingly rare as you know. So the wrong implant and giving the wrong, like a so-called sentinel event, the rates on that are minuscule. It's a lot because there are a million cataract surgeries or whatever, but the rate is minuscule. And as far as worrying about operating on people better than 20/40, I'm not a cataract surgeon, but there are Medicare exclusions. You can't get paid for operating on people who are better than 20/40 unless there are very special things met. So there aren't many of those.

MEMBER STEIN: I'm the other lead on this one. I think in terms of the evidence, it's clear that improving someone's vision with cataract surgery is a good thing. It improves quality of life, it improves ability to be productive, ability to drive. So I think the issue with the evidence, let's not bring in some of these other issues with the measure and just

focus on the evidence first.

CO-CHAIR YAREMCHUK: Okay. So, can we do a vote on this?

MEMBER CARNAHAN: We're saying the evidence here shows the patient had an improved vision, even though we don't know that to be true?

CO-CHAIR YAREMCHUK: That they're not any worse.

DR. WINKLER: The evidence criteria, if you recall is that there is a rationale that supports the relationship of this outcome with some healthcare process service intervention. So in other words is it actionable? Can you affect the outcome? Is there something that anybody can do that could that could affect the outcome? That's all you need. It's basic for evidence for an outcome measure.

DR. RICH: I'll point out that when you look at large numbers like we are now, there are huge variability between surgeons on this measure. Again, this is a patient-focused

measure with public reporting.

CO-CHAIR YAREMCHUK: So, just to follow up, because of the criteria difference, your voting options are yes and no. There is the rationale.

MS. ROBINSON-ECTOR: So, voting's now open for evidence for measure 0565, and for those on the call, option 1 is yes, and option 2 is no. Oh, it's up here. I don't know why it's not there, but it's up here. I can see it. Okay so all votes are in and voting is now closed. Sorry, I can see it in not private, yes. So 93 percent voted yes, and 7 percent voted no. So, for evidence, measure 0565 passes.

MS. LUONG: And this is with 15 votes from the committee?

MS. ROBINSON-ECTOR: Yes.

CO-CHAIR YAREMCHUK: Opportunity for improvement?

MEMBER CARNAHAN: Sure. So, there's a lot of data out there. The PQRS data, again, small numbers in voluntary reporting and so

forth, 90.6 percent in 2010, 94.8 in 2011, in 2012 92 percent. So it's definitely not a trend that's been showing success in and of itself. So although I guess the flip side is, does that mean the measure is not effecting positive change? We don't know. The CRS had some data, 85.5 percent, so it kind of falls in that range. And with over three million cataracts per year across the U.S., that's hundreds of thousands of patients. That's a lot of potential for improvement and a lot of patients that are impacted.

CO-CHAIR YAREMCHUK: Anything to add?

MEMBER STEIN: No, I think that's well stated.

MS. ROBINSON-ECTOR: So voting is open for performance gap for measure 0565 and for those on the call, option 1 is high, 2 is moderate, 3 is low, and 4 is insufficient. Okay, all the votes are in. 60 percent voted high, 40 percent voted moderate, zero voted low, and zero voted insufficient. So performance gap for measure 0565, the measure passes.

CO-CHAIR YAREMCHUK: Okay, reliability?

DR. WINKLER: Now just--for right now, I'd really like to focus on the reliability of measure in the registry. Just in the registry. We'll talk about the eMeasure version afterwards. Okay? So, right now, focusing in on the information around the reliability of the measure as a registry measure.

MEMBER CARNAHAN: So with clear-- well I guess the big thing is the exclusions, and there's a long list of exclusions, but I would say almost all that have the potential to cause a vision reduction in and of themselves. So if you didn't have those exclusions, then you would have a significant reduction in the results and I think that when looking at the exclusions, it led to about a 50 percent reduction in the volume of surgeries that we're counting. Sorry. Anyone hear any of that? So, there's a long list of exclusions and they are all vision threatening exclusions or vision reducing exclusions above

and beyond the cataract itself. So as Dr. Rich said, it keeps the data more clean to focus on just where the opportunity is around the cataracts. It would be a separate measure if it included all the exclusions. With the exclusions not being counted, it reduced the volume to about half and with the results we've already talked about.

MEMBER STEIN: I had a few issues with reliability, none of which I think are major issues, but things for the developers to think about. Matt brought up the point about patients who start off with better than 20/40 or better vision. And could someone game the system by operating earlier to improve their chances of being successful? It might be good in a future iteration to exclude those from the denominator. The issue of when the measurements are being made is another--if someone's coming in multiple times at what visit they're to get the--if someone comes in ten times, their chances of getting 20/40 or better vision is better than someone who

only comes for one or two post-op visits; so incorporating that in.

I'm not sure how good the registry--I think that fine tuning how the vision is being captured in terms of manifest refraction, best corrective visual acuity, different providers may be capturing different things, so just to make sure that you're capturing the same thing and that provider A is being compared against the same metric as provider B. And then, it wasn't clear why, if we have all these exclusions, why not also exclude people with a more severe cataract, the 66982 as another exclusion and just focus on the people with the 66984?

DR. RICH: Great questions. There's three things that you addressed. Number one is gaming operate on someone with less than 20/40. There's a word for that, it's called fraud. That will enable you to visit one of those long states in the middle of the country. And as far as when we look at the registry, we can actually map out when the vision is measured, and I think the

average number of visits is about 3.5. Flora?

DR. LUM: Right. It's within that 90 day period.

DR. RICH: In that 90 day period. What was the third one, Josh?

MEMBER STEIN: Why not exclude 66982?

DR. RICH: Oh, because there's no underlying disease that precludes you, that's a complicated cataract. That's one that's tough. We thought that's a measure of surgical competency. There's no underlying co-morbidities that would prevent you from seeing 20/40. So we felt it was important to leave that in.

MEMBER FRIEDMAN: So, a couple more comments. Measuring vision is somewhat subjective. For those of us that--there's ways to specifically quantify vision and be certified, et cetera. We don't do that. This is the best way to measure vision given the limitations of measuring vision in general. The reason we exclude a lot of patients is there's a lot of people out there that don't have a vision

potential of 20/40, but clearly benefit by cataract surgery. So there are a lot of people that are only going to see 20/60, but this is going to be their better seeing eye than the other eye that can only see light and dark. And to get them from 20/200 to see the big E, to seeing 20/60, is markedly going to improve their quality of life. They're not going to be able to drive in most states, but they will be able to read with better reliability and they're going to be happier and the quality is going to improve. But that's why there's so many exclusions for this measure.

MEMBER MADONNA: Just getting back to what Josh said about multiple visits and you said 3.5 is the average, but isn't it possible then that somebody comes in at day eight and has 20/30 vision, then develops cystoid macular edema and at the end of 90 days has 20/100 and perhaps they never recover 20/40 vision again. So they would pass this, but yet that's going to be a very unhappy patient.

DR. RICH: Well you know, that would be a failure if your developed CME and I guess the measure does say within 90 days.

MEMBER BRADHAM: This is Tammy and I'm coming at this as a parent with a daughter with significant vision issues. And when I go in for those evaluations, they always test monocular and then both of the eyes together. And when I read this definition it doesn't come out and say--it says cataract surgery, but are you getting at the vision tests with both eyes or just one eye? It's not---

DR. RICH: Just the surgical eye.

MEMBER BRADHAM: Right, but it's not clear in the description.

DR RICH: I don't think, you know, we--once we're in that 90 day period, we're not even checking the other eye.

MEMBER BRADHAM: I know, but from a parent--if this is written for the consumer, it's not clear. I'm just saying that.

MEMBER CARNAHAN: It does state though

that each eye is a separate episode, and so I think that's how they're saying, they're clarifying that piece of the---

CO-CHAIR YAREMCHUK: Steve?

MEMBER GOLDBERG: Does the development of cystoid macular degeneration following cataract surgery though related to how it's performed, or is it an unavoidable complication?

MEMBER FRIEDMAN: I think it can be, certainly. If someone's spending more time taking the cataract out or has poor surgical technique, I think it's possible certainly to have a greater risk of cystoid macular edema.

MEMBER SCHACHAT: So the answer to that is it's some of both? Overall the rate of visually significant cystoid macular edema over the past few days, ends up being one percent. About thirty percent of patients have some of it but it doesn't affect their vision, but affecting the vision is about one percent. And some of it relates to messy, botched, longer, more tedious surgery, and a lot of it relates to other

conditions that got facilitated. So it's probably about half and half, but the overall rates are low.

I just wanted to make a comment about the concern about the multiple visits in 90 days and what Josh was trying to get at is should the vision be specified as measuring at a certain time point rather than getting a positive score for any time in the 90 day period, and then the whole thing wouldn't work because there's too much variation on the timing of visits. If you want to measure at 90 days, so many patients don't have to come back at 90 days, some would have to come back from Chicago to Cleveland to get measured at that 90 day visit approximately when they don't actually need it. So I think there's going to be a little mess in variation in the measure that I would accept to allow patients to have flexibility on their visit time.

MEMBER STEIN: Yes I don't think any of these issues I brought up are critical, major issues. It's just things for the developer to

think about to enhance the tool they have.

CO-CHAIR YAREMCHUK: Any other questions?

MEMBER FRIEDMAN: One more quick comment. So just getting back to the CME. Basically, CME is very rare and it's usually treatable. It usually goes away but if there's a provider that has, who's an outlier that has a significant number of his patients developing CME, that's a problem and that's what we want to try and identify.

CO-CHAIR YAREMCHUK: So, are we ready to--oh excuse me. Go ahead.

MEMBER CARNAHAN: Just back to the first point that it seems like the measures that you're working on to say how close are you to your refractive goal and maybe did you have an improvement in vision after cataract surgery. Because even the patient who is not fraud, who is a 20/40 cataract and comes out 20/40 who you think theoretically he'd be seeing better afterwards, is going to be a success here even

though they saw no better after surgery. But if the measure was saying an improvement over pre-operative vision, or met the target within half a diopter or whatever the new measures are you're working on, I think that would be really clean. And so perhaps once those kind of measures come along, then this measure falls to the wayside.

MEMBER FRIEDMAN: So again, for those of us who do clinical research, there's different ways to measure visual function. One way is known as acuity, which is what 20/40 means. You can also look --- there are people with 20/40 that have a lot of glare. They can't drive at night, for example. So they go from 20/40 to not being able to drive at night, to 20/40 being able to drive at night; their quality improves markedly. We're not measuring that actually, but so there is an indication for people that are 20/40 to have cataract surgery in lieu of not actually having a visual improvement or improvement in their -- acuity, but their visual function does in fact improve.

DR. RICH: There's a separate independent measure that --- and you know, I don't want to prolong this discussion, but looks at the patient reported outcome pre and post op, separate from this.

CO-CHAIR YAREMCHUK: Right. Are we ready to vote on reliability?

MS. ROBINSON-ECTOR: Voting for reliability for measure 0565 is now open and for those on the line, option 1 is high, option 2 is moderate, option 3 is low, and 4 is insufficient. All the votes are in. 13 percent voted high, 80 percent voted moderate, 7 percent voted low and zero voted insufficient. So for reliability, measure 0565 passes.

CO-CHAIR YAREMCHUK: Okay. Validity?

MEMBER CARNAHAN: So for validity, to report on when it was tested. It was looked at in a single practice and I think that was an area of contention for some people because it was three physicians and one EHR, but it had a 92.6 percent agreement. The PQRS was evaluated with

454 physicians, with 403 having enough data to submit for this, and the only trouble with that was it didn't identify the exclusion piece of things. And as we talked about before, the exclusions, the potential threats were that 47.8 percent of data remained after all the exclusions were included. And I guess the missing data again, I would say would be a pre-operative comparison or a vision target.

And, again, back to Scott's comments, as long as you had an improvement in vision, then you would --- I don't know how you'd quantify that, 20/40 can be different kinds of 20/40, but I would say that now you get a larger pool of patients. You get all the glaucoma patients and macular degeneration patients who go from 20/400 to 20/80 and they're still a success. And so you have done them a good as opposed to their being excluded in these kinds of situations. And so they may be put to the wayside. Hopefully not.

CO-CHAIR YAREMCHUK: And can you just talk up a little bit. Josh, do you have

anything?

MEMBER STEIN: The other validity test they did was the face validity with an expert panel of 16 members and found good validity that way.

CO-CHAIR MERENSTEIN: I'm just going to need all of you to speak up. They can't hear us back there.

CO-CHAIR YAREMCHUK: So, any comments? Ready to vote on validity?

MS. ROBINSON-ECTOR: Voting is now open for validity for measure 0565 and for those on the line, option 1 is high, 2 is moderate, 3 is low, and 4 is insufficient. Okay, all the votes are in. 20 percent voted high, 80 percent voted moderate, zero voted low, and zero voted insufficient. So for validity, measure 0565 passes.

CO-CHAIR YAREMCHUK: Okay. Feasibility?

MEMBER CARNAHAN: This has a very specific category, two codes that would identify

a best corrected vision of 20/40 or better, distance or near, achieved within 90 days and a separate code if it was not achieved, so it would be very feasible that this would be easy to pull the data on.

MEMBER STEIN: No, I think that Matt summed it well.

CO-CHAIR YAREMCHUK: Okay. Ready to vote on feasibility?

MS. ROBINSON-ECTOR: Voting for feasibility for measure 0565 is now open and for those on the call, option 1 is high, 2 is moderate, 3 is low, and 4 is insufficient. All the votes are in. 60 percent voted high, 40 percent voted moderate, zero voted low, and zero voted insufficient. So, for feasibility, measure 0565 passes.

CO-CHAIR YAREMCHUK: And the last metric is usability and use.

MEMBER CARNAHAN: As we talked about, there's PQRS data for this. It goes to IRIS registry and publicly reported as a result. We

talked about the room for improvement with it being in the low 90's right now and that being hundreds of thousands of patients concerning the high volume. And I think just the only unintended consequences would be it might falsely identify a good surgeon who wasn't because there was no vision change potentially. As is also mentioned, it may end up increasing return visits so that you can get a successful outcome by just bringing them back until you finally get a 20/40, more than you would have otherwise.

CO-CHAIR YAREMCHUK: Comments?

MEMBER STEIN: The only other unintended consequence would be a surgeon not wanting to operate on someone that they didn't think could get to 20/40, but hopefully that's not going to happen. Hopefully, that wouldn't happen that much. I don't think that's a major issue.

CO-CHAIR YAREMCHUK: Any other comments? Then we can vote.

MS. ROBINSON-ECTOR: Voting for

usability and use for measure 0565 is now open and for those on the line, option 1 is high, 2 is moderate, 3 is low, and 4 is insufficient information. All the votes are in. 60 percent voted high, 40 percent voted moderate, zero voted low and zero voted insufficient information. So for usability and use for measure 0565, the measure passes.

CO-CHAIR YAREMCHUK: Okay. Last vote. Whether to recommend measure as suitable for endorsement. Any comments from anyone?

MS. LUONG: And this is for just clarification purposes, the registry version of measure 0565.

MS. ROBINSON-ECTOR: The vote is now open for recommendation for overall suitability for endorsement for measure 0565 for those on the call, option 1 is yes, and option 2 is no. All the votes are in. 100 percent voted yes and zero voted no. So for recommendation for overall suitability for endorsement for measure 0565, measure passes.

MS. LUONG: Again, to note that is for the registry version.

CO-CHAIR YAREMCHUK: All right, go ahead.

DR. WINKLER: As I mentioned --- go to the next slide --- as I mentioned, this measure does have an eMeasure version, in the way we discussed prior. So the issue around evidence and gap shouldn't be any different, so we don't need to revisit that. So we do want to talk about the potential concerns or issues or questions around the reliability and validity, the scientific acceptability of the eMeasure. Because this measure is one of the original retooled measures, they haven't -- and as I talked about earlier, this measure hasn't been tested widely. We do have limited testing data in one EHR system. We've already talked with the developers, they're willing to provide us the information for the simulated data set. And that sounds like something that can come back soon-ish, like the next time we meet, and so you could

see that. And that would be acceptable at this point in time.

We're certainly hoping that as more and more eMeasures are used and we're collecting data, we can actually do better reliability and validity statistical analysis on the data generated from eMeasures, but at this point in time, it's pretty much a limitation across the field. So we're just living in a transition time. So, what I'd like to do is just have the group talk about whether you believe that what information we have on reliability and validity, it's a single vote, for the eMeasure is acceptable with the conditions we've talked about. And then --- so Kaitlynn's going to put that one up. So does anybody have any comments about the eMeasure version from the committee first? Flora, I'll give you a chance.

DR. LUM: Brief comments from the developer. This was discussed before. We have extracted it for 2,722 physicians in 2014 across 26 different EHR systems. And all our files were

accepted successfully by CMS. So we haven't had any difficulty in extracting the data. And in addition, we review the rates with the physicians, and ophthalmologists are very compulsive, so if patients are failing the measure, they can look back on which patients are failing the measure and we look back to look at the measure to see if there's any problems with the calculations. So, they all approve of what we've submitted on behalf of the --

DR. WINKLER: So I am -- as I said, because we are considering it, I am going to ask you to vote, we're going to combine reliability and validity into one question. It's a yes/no question. Is the information that we have about eMeasure acceptable for scientific acceptability with the conditions we're going to get the additional information from the simulated data set?

MS. ROBINSON-ECTOR: So voting is now open for reliability and validity for the e-version measure of 0565 and for those on the

call, option 1 is yes with conditions, and option 2 is no. All the votes are in. 100 percent voted yes with conditions and zero percent voted no. So for reliability and validity, the eMeasure version of 0565 passes.

DR. WINKLER: So when it comes to feasibility, there are often feasibility issues or concerns raised in the discussion about use of an eMeasure or any potential way the measure is specified in terms of data elements, whether these are routinely collected and whether -- so those are the considerations. I think Flora's provided you some information about their experience. But in terms of your assessment of feasibility of the actual use of this particular eMeasure in EHR systems is the question for feasibility for the eMeasure. Questions, comments? Okay, it sounds like you're ready to vote.

MS. ROBINSON-ECTOR: Voting is now open for feasibility for the e-version measure of 0565 and those on the call, option 1 is high, 2

is moderate, 3 is low and 4 is insufficient. All the votes are in. For feasibility, 60 percent voted high and 40 percent voted moderate, zero voted low and zero voted insufficient. So for feasibility for the eMeasure version of 0565, the measure passes.

DR. WINKLER: Okay, now for usability and use, this is one where perhaps your previous conversation would apply. Is there anything different about the eMeasure or would you like to just transfer your votes from the prior --- from the non-eMeasure version and have it apply here as well? Does anybody have any --- does that seem accurate and a reflection of your thoughts? Okay. Then we'll stipulate the use and usability and then the last one will be the recommendation for endorsement with the conditions we get the follow-up information.

MS. ROBINSON-ECTOR: Oh sorry. So this is the side that we'll be voting on. Voting is now open for overall suitability for endorsement for eMeasures for measure 0565 and

for those on the call, option 1 is yes with conditions and option 2 is no. All the votes are in and voting is now closed. 100 percent voted yes with conditions and zero percent voted no. So for overall --- recommendations for overall suitability for endorsement for the e-version measure of 0565, the measure passes.

DR. WINKLER: Just for the transcript record, the condition is that they will bring back the testing from the simulated data set using the Bonnie tool at the next time we meet after the comment period for you to evaluate.

CO-CHAIR YAREMCHUK: We were supposed to have break at 2:30 and it's 3:08. So yes, it's break. The question is when do we reconvene? 20 after? Okay, 20 after, reconvene.

MEMBER LYNCH: This is Judith, I'm going to have to sign off for today; I'll be back on in the morning.

[Whereupon, a break was taken from 3:02 p.m. to 3:16 p.m.]

CO-CHAIR YAREMCHUK: Okay. I'm hoping to gather

everybody back around the table. I have two requests. I guess I'm going to call these housekeeping requests.

One is to speak in the microphone. And speak loudly enough. One, the people behind us are having difficulty hearing.

And there's also someone transcribing. And so it's difficult for them to be able to transcribe. So, to make a conscious effort to do that.

And we can now go to the -- is everyone back? Yes? I think so. Okay.

So then the next measure for consideration is 0564, Complications Within 30 Days Following Cataract Surgery Requiring Additional Surgical Procedures. And so Matt, you're up again.

MEMBER CARNAHAN: All right. So, these again are the same population of people, 18 years and older. Have an uncomplicated cataract. And there's a long list of those. And I guess we'll talk about that later on for the exclusions

piece.

Who had one of a number of situations arise after surgery. Retain nuclear fragments and endophthalmitis, dislocated or wrong power interocular lens, wound dehiscence.

But the one question that came up for our group was retinal detachment. Meaning that you could definitely induce a retinal detachment during surgery.

But at the same time, I would venture that most of the retinal detachments that happen after surgery are probably in patients who may have had them anyway.

And so the question is, are we identifying anything around a less safe surgeon who, perhaps, has a perfect surgery and yet, has a patient with retinal detachment? So --

MEMBER SCHACHAT: So, the rate of retinal detachment after cataract surgery needs a time denominator. It's ever increasing and it goes up forever.

And it has to do mainly with the time

of their developing PVDs, posterior vitreous detachment for the retinal -- for the ENT doctors.

And those that occur within 30 days are a very small fraction of the overall number that are going to occur. And I think they're more likely to relate to surgical mess ups, as opposed to just the routine kind of retinal detachment.

It's going to happen from PVDs, which happens at a regular rate forever. I think.

MEMBER CARNAHAN: And that goes along with I think another measure they have around posterior capsular rupture, which would definitely increase your rate of retinal detachment.

CO-CHAIR YAREMCHUK: Okay.

MEMBER CARNAHAN: So, just putting that out there for the evidence piece. Exclusions we talk about later, correct? Okay.

So we're just trying to identify complications that can reasonably be attributed

to the surgery or the surgeon. And reflect situations that if untreated, generally result in avoidable vision loss.

CO-CHAIRMAN YAREMCHUK: And once again, this is an outcome measure. And so what I had read before report, whether there are process of care that can influence the outcome.

Nothing more is required about evidence for outcome measures. Okay. Do we need to vote on this?

MS. ROBINSON-ECTOR: Yes.

CO-CHAIR YAREMCHUK: Okay.

MS. ROBINSON-ECTOR: Okay. So voting for evidence is now open for Measure 0564. And for those on the line, option one is yes. And option two is no.

MS. LUONG: And to note, this is not for the E-version of the measure.

DR. WINKLER: It applies to both.

MS. ROBINSON-ECTOR: And we're just waiting for the calls -- the people on the line to vote. Yes, so we're just waiting for one.

Okay. All the votes are in. And voting is now closed. 93 percent voted yes. And seven percent voted no. So for the evidence of Measure 0564, the measure passes.

CO-CHAIR YAREMCHUK: All right. Do you want to discuss opportunity for improvement?

MEMBER CARNAHAN: Sure. So looking at the PQRS data from 2010 through 2012, there's actually been a higher percentage from 3.4 to 4.4 to 5.2, which would indicate a lack of improvement. And therefore, a greater opportunity for improvement over time.

And looking at the registry numbers, it goes from zero to one in terms of with a median of zero. But still, there are those with a high opportunity for improvement.

So, I would say that there's definitely room considering the huge three million plus population of patients we're talking about.

MEMBER STEIN: Yes, I think given that cataract surgery is the most common surgery there

is, you know, three to five percent means that there are many patients experiencing these issues.

You know, the fact the rates are going up over time, it's a little hard to interpret whether that is -- that people are having more complications or that our coding of complications is getting better with EMRs and other captures.

DR. RICH: I think it's a small sample size, Josh. A very, very small sample.

MEMBER STEIN: Right.

MEMBER FRIEDMAN: I think just as Bill suggested, you have to look at the confidence intervals around those numbers. If the confidence intervals are very large, they could be not -- they could just be noise and not meaningful differences.

CO-CHAIR YAREMCHUK: All right. Are we ready to vote?

MS. ROBINSON-ECTOR: Voting is open for performance gap for Measure 0564. And for those on the line, option one is high, two is

moderate, three is low and four is insufficient.

All the votes are in and voting is now closed. 36 percent voted high. 64 percent voted moderate. Zero voted low and zero voted insufficient.

So for performance gap, Measure 0564 passes.

CO-CHAIR YAREMCHUK: Okay. Validity? Oh, sorry, reliability?

MEMBER CARNAHAN: So, looking at the reliability, the numerator statement has a list of certain procedures and a certain diagnosis. And then there's like I said, a large list of exclusions.

I mentioned the retinal detachment piece already. But in terms of the exclusions, it includes things like mature cataracts, senile cataract, glaucoma.

I would wonder if those were all necessary exclusions. And then Alpha-1 blocker use, definitely it can be a problem. But I would imagine that at least 15 percent of the time it

wouldn't be.

So, I'm just wondering if we're over-excluding, any -- just the easiest of the easy cataracts that people can therefore be successful to measure, as opposed to truly identifying those who may need some help?

CO-CHAIR YAREMCHUK: Go ahead. Josh?

MEMBER STEIN: This came up with the previous measure, too. That I don't think it's necessarily a fundamental problem with this measure.

But I'd like to see in subsequent measures to expand the number eligible. You know, have one measure for complete uncomplicated cataracts. And then another for cataracts with issues.

And hopefully, the developers are looking to do that sort of thing.

DR. RICH: We're doing exactly that. We'll be able to look at people with pseudoexfoliation. You know, all those issues we're actually collecting massive amounts of data

now.

We'll have over a million next year. So we'll be able to stratify those.

DR. LUM: I just wanted to add that I know the analysis you were provided looked at Medicare claims data over 70,000 procedures. But the IRIS Registry is the ECQM, which is all patient populations.

So, the claims were limited to Medicare Part B, only fee for service. So we had over 300,000 procedures.

And the exclusion was much lower. It was 27 percent for that measure. So, we are including many more patients.

And as Dr. Rich said, we provide the rate for both complicated and uncomplicated patients to our participants, so that they can look at improvement in quality and look at those patients that had a complication. And really dig down and see what was causing it.

And it's really for their own quality improvement purposes.

CO-CHAIR YAREMCHUK: Steve?

MEMBER STRODE: The previous measure was at 90 days. And this is at 30 days. And my question is, would 30 days capture most of those complication surgeries?

Recognizing that the timing of such surgery goes beyond the recommendation of the ophthalmologist. It might have patient issues. It might have health insurance approval issues and others.

DR. RICH: The vast majority of the return to the OR is going to occur from surgical error. And statistically, all most all of them are somewhere around ten days or less.

Or certainly up to 30 days. In 90 days, you might be picking up a couple of people that had an incidental retinal detachment.

So we felt that it was -- we could actually capture most of those surgical errors that led to the complications and we return to the OR. They can be captured both in EHR and the claims data within 30 days.

CO-CHAIR YAREMCHUK: Are we ready to vote?

MS. ROBINSON-ECTOR: Voting is now open for reliability for the registry version of Measure 0564. And for those on the line, option one is high, two is moderate, three is low and four is insufficient.

I'm just waiting on one more vote.

MS. LUONG: Vaishali, if you can just email -- thanks, got it.

MS. ROBINSON-ECTOR: Great. All the votes are in. 43 percent voted high. 57 percent voted moderate. Zero voted low and zero voted insufficient.

So for reliability, the registry version of Measure 0564 passes.

CO-CHAIR YAREMCHUK: Okay. Can we go on to validity?

MEMBER CARNAHAN: This is very similar to the prior measure where the expert panelists 16 strongly agreed the measure could be distinguished as a quality of care measure.

There's also this small claim -- the small group of one practice, one EHR that was evaluated with a 99 to 100 percent agreement. Then there's the registry patients with a high level of agreement.

So, we already discussed the exclusions piece of things. And then there was some socioeconomic questions as to if that was taken into account. And maybe that would be something the IRIS Registry would be able to stratify as well in the future.

MEMBER STEIN: Yes, I don't think we're ready for risk adjustment yet. But I'd like to see it in the next round when this comes up.

Risk adjustment as clearly, for example, Asian Americans have more myopia and are more prone to retinal detachments in some of these issues. So, I think in the future that would be something important to incorporate.

DR. RICH: Exactly what we're doing.

CO-CHAIR YAREMCHUK: All right, are we

ready to vote?

MS. ROBINSON-ECTOR: Yes. The voting is now open for validity for Measure 0564, for the registry version of the measure.

Oh, sorry. And for those on the line, option one is high, two is moderate, three is low and four is insufficient.

All the votes are in. 21 percent voted high. 79 percent voted moderate. Zero voted low and zero voted insufficient. So for the registry version of Measure 0564, the measure passes validity.

CO-CHAIR YAREMCHUK: Usability and use?

MEMBER CARNAHAN: Feasibility? So, just like the prior Measure, very feasible. All the diagnosis as well as all the procedures have distinct codes that can be easily data mined.

So, I guess if there was -- it says, indicate whether any feasibility concerns such as fees for registry, participation, I suppose there is, as you would have to have an AAO membership

to participate in the IRIS Program.

But, otherwise you can do it through the PQRS. And that should be easily mined from any EHR I would think.

CO-CHAIR YAREMCHUK: Josh?

MEMBER STEIN: Nothing to add.

CO-CHAIR YAREMCHUK: Okay. Vote?

MS. ROBINSON-ECTOR: Voting is now open for feasibility for the registry version of Measure 0564. And for those on the line, option one is high, two is moderate, three is low and four is insufficient.

All the votes are in. 64 percent voted high. 36 percent voted moderate. Zero voted low and zero voted insufficient.

So for feasibility, for the registry version of Measure 0564, the measure passes.

CO-CHAIR YAREMCHUK: Okay. Now, usability and use.

MEMBER CARNAHAN: As we talked about the PQRS options, the IRIS options, there's the meaningful use component of things. So this

should be very easy, usability and use would minimally impact any practice.

MEMBER STEIN: Nothing to add.

CO-CHAIR YAREMCHUK: Vote?

MS. ROBINSON-ECTOR: Voting is now open for usability and use of the registry version of Measure 0564. And for those on the call, option one is high, two is moderate, three is low and four is insufficient information.

Okay. All the votes are in and voting is now closed. 50 percent voted high. 43 percent voted moderate. Zero voted low and zero voted insufficient information.

So for usability and use of the registry version of 0564, the measure passes.

CO-CHAIR YAREMCHUK: Okay. Should we vote on E-measures? Or that's after we do the --

MS. ROBINSON-ECTOR: After we do this.

CO-CHAIR YAREMCHUK: Okay. So we'll vote on endorsement now for the measure.

MEMBER CARNAHAN: Let me make one more comment. Just that the panel had suggested that

the title actually should be selective complications as opposed to complications. Because it would indicate that someone wasn't having other complications.

And yet, I think this is, perhaps, the minority of complications that happen. I think a posterior capsular rupture is much more common than any of these things combined.

Another piece of it is that you have to have both hits to have a trigger. And I would venture to say that over half the time, and I think I'm being generous, a patient who has a wrong lens implanted is not going to go back and have a lens exchanged.

They're going to have laser vision correction procedure that's not going to trigger this. Just because that's the way the world works.

So, I don't think you can fix that. I don't think that's possible. But I just think that's reality.

DR. RICH: Yes. I think in the --

actually out in the community, people would exchange that lens, rather than taking them back for a second procedure.

MEMBER CARNAHAN: Are they just wearing it?

DR. RICH: Yes.

MEMBER CARNAHAN: On a laser and they put a multi-focal lens in and it's not quite right, they're not going to take it out and put another one in. They're going to do the PRK which costs them $350, versus losing money and doing a whole new cataract.

And that seems like it should be recentered.

DR. RICH: So I think we could -- the afternoon is getting late, but --- I think I'll let that one pass. The whole idea of having selective complications is you could actually quantify them with claims data.

So these are things that you can't fudge. There's going to be a return to the operating room.

And that's why we have a separate measure for vitreous loss, specifically. But, see that's much harder to track. And -- but these are much easier to track.

They're made more -- and they're a little more serious things --- wrong eye well, dropped nucleus, all the things that Scott sees in his office after.

MEMBER FRIEDMAN: I don't see any of those. So, one more question. I just started thinking, how do you differentiate returning to the operating room that isn't related to the complications?

So the patient gets surgery. Then they go to the operating room within 30 days, but it's not a complication of the cataract surgery?

MR. RICH: On an eye procedure?

MEMBER FRIEDMAN: Well, yes. Diabetic macular edema.

MR. RICH: I don't know that I can answer that. Flora?

MEMBER SCHACHAT: I think the safety

catch on your concern is that most or probably the overall majority relate to the surgery. I mean, the one that I would worry about is stuff on the other eye.

But there are eye modifiers that let us crack that.

MEMBER CARNAHAN: And it does have very specific diagnosis for the returns. Not any return to the OR in 30 days.

It's removal procedures, excision of adhesions. Tap and inject for the ophthalmitis one lens procedures. Retinal repair, but not necessarily a vitrectomy.

So it doesn't go through every possible return to the OR. But only selective ones I imagine they felt would be related to the very select diagnoses of that.

MEMBER STEIN: The use of modifiers should be able to distinguish that. You know, it's conceivable that someone could have such a dense cataract that you can't see what's going on behind it.

And then you take it out and realize there are other issues that require additional surgery. But the modifiers should sort that out.

CO-CHAIR YAREMCHUK: Okay. Are we ready to vote on endorsement? Okay.

MS. ROBINSON-ECTOR: So, the voting is now open for recommendation for overall suitability for endorsement for the registry version of Measure 0564. And for those on the call, option one is yes and option two is no.

Okay. All the votes are in. 100 percent voted yes. And zero percent voted no. So for recommendation for overall suitability for endorsement for the registry version of Measure 0564, the measure passes.

DR. WINKLER: All right. This measure also has an eMeasure version. So as we did with the last measure, we do want to look at any particular issues that might be specific to the eMeasure.

That evidence and gap should be the same as for the other. So we're talking about

scientific acceptability.

We have the same issues that they will be doing the additional testing in a simulated data set and providing us the information from the results of that testing when we regroup after the comment call.

And so, the voting is as before for scientific acceptability of the eMeasure, you know, with those conditions.

Does anyone have any comments or questions about the eMeasure specifically in terms of reliability and validity? That may be distinct from the registry measure? Okay.

Ready to vote on measures. Yes or no, and once again, yes is with conditions. And the condition is bringing back the information from the testing of the simulated data set.

MS. LUONG: And this is for the eMeasure version of 0564. So Vaishali, one is yes with conditions and two is no.

MEMBER PATEL: Oh, sorry. I thought I already voted. Sorry.

MS. ROBINSON-ECTOR: Great. All the votes are in. Okay. For reliability and validity, for the eMeasure version of 0564, 100 percent voted yes. And zero voted no. so the measure passes.

CO-CHAIR YAREMCHUK: Next up is feasibility of the eMeasure. Does anyone have any questions, concerns, comments about feasibility of this as an eMeasure?

Okay. Then we'll ask you to vote on that.

MS. ROBINSON-ECTOR: Voting is now open for feasibility for Measure 0564, for the eMeasure version. And for those on the line, option one is high, two is moderate, three is low and four is insufficient.

Okay. All the votes are in. 64 percent voted high, 36 percent voted moderate. Zero voted low and zero voted insufficient. So for feasibility for the eMeasure version of 0564, the Measure passes.

DR. WINKLER: Now, as we did before,

is there anything different or distinct about the eMeasure as to usability and use that you would want to vote it separately? Or should we just use the same voting you did for the registry version for usability and use?

Anybody have any objection to using the same results? Excellent.

So we can move on to the recommendation for endorsement. Again, this is the -- with the conditions of the additional information that will be brought back around the simulated data set.

MS. ROBINSON-ECTOR: So, voting is now open for recommendation for overall suitability for endorsement for the eMeasure version of 0564. And for those on the call, option one is yes with conditions and option two is no.

All the votes are in. 100 percent voted yes with conditions. And zero voted no. So for recommendation for overall suitability for endorsement for the E-Measure version of 0564, the measure passes.

CO-CHAIR YAREMCHUK: So we are next -- for our next measure, is this also an eMeasure?

DR. WINKLER: No, it's not.

CO-CHAIR YAREMCHUK: Okay. So 0563, Primary Open-Angle Glaucoma: Reduction of Interocular Pressure by 15 Percent or Documentation of a Plan of Care, American Academy of Opthalmology. And who is going to discuss this? Okay.

MEMBER STEIN: So we're discussing the evidence now?

CO-CHAIR YAREMCHUK: Correct.

MEMBER STEIN: Okay. So there are several large landmark randomized clinical trials that have shown that lowering eye pressure can prevent worsening of vision and blindness from glaucoma.

The American Academy, as a preferred practice pattern that recommends lowering of eye pressure of 25 percent or more. In the clinical trials that are out there, show levels of 18 to 42 percent.

And it's a Grade A, Level Two evidence. So I think there's evidence to support the measure.

CO-CHAIR YAREMCHUK: Is there someone that's your assistant or is also on this?

MEMBER PATEL: Well, yes. Vaishali Patel on the phone. So I'm the other discussant for this measure.

So, I agree with Josh that there's enough evidence. Plenty of enough evidence to show importance of lowering IOP for treatment of glaucoma.

And it's the only modifier of this factor. And so from that perspective, it's, you know, an important thing to do than to lower.

CO-CHAIR YAREMCHUK: Okay.

MEMBER PATEL: ---- and to measure.

CO-CHAIR YAREMCHUK: Any comments? All right we can --

MEMBER STEIN: There's a comment.

DR. RICH: Yes, just a point of information that was discussed in the call was

the plan of care. And we felt it was very important to have the failure in there.

And if you look at that, that's actually consistent with all of our other Measures. The hemoglobin A1c measure, the hypertension measures.

There's a certain number of people where you make the value judgment that you're going to have more complications. And so you're willing to follow them.

So, that's why that plan of care is in there. And we were asked to document how many people passed the measure by looking at a plan of care when they didn't get there.

And we looked very carefully and the number was 30 percent. So, that's actually kind of consistent with the -- our other chronic disease measures that have a goal in mind.

CO-CHAIR YAREMCHUK: Okay. Are we ready to vote? Evidence?

MS. ROBINSON-ECTOR: Voting is now open for evidence for Measure 0563. And for

those on the call, option one is high, two is moderate, three is low and four is insufficient evidence.

Okay. All the votes are in. 64 percent voted high. 36 percent voted moderate. Zero voted low and zero voted insufficient. So, for evidence of 0563, the measure passes.

CO-CHAIR YAREMCHUK: Okay. Opportunities for improvement?

MEMBER STEIN: So, according to the information provided by the developer, the mean performance from 2009 to 2012 is 93 to 95 percent. So there is some room for improvement.

DR. LUM: I'd just like to add that was -- that number at least in 2013 only represented 15 percent of the eligible providers. So we're still talking about a very small minority of patients.

The other thing I wanted to bring up I think is we're really going through a transformational demographic trend. And we haven't talked about disparities in care.

But I think this issue, and we'll get to diabetic retinopathy, are huge issues for us in eye care. And just for the non-eye care professionals, I just wanted to explain that we're going from 2.71 million people with POAG to 7.31 million in 2050.

And the largest demographic group is changing from older white women to Hispanic men. And that's by 2035. We're going to double per capita rates of POAG in key states like New Mexico, Texas and Florida because of the demographics.

And we already know that Hispanics have less access to care. In fact, Dr. Stein wrote an article showing the odds of ancillary testing for glaucoma are less in Hispanics.

So I do want to keep that focus on these population groups that potentially have less access to care, are in rural areas, under-served areas and potentially with providers who are less knowledgeable.

CO-CHAIR YAREMCHUK: And I guess the

second part is this -- it says briefly discuss any data on disparities.

DR. RICH: Yes. And then African Americans have a threefold higher rate of blindness and incidents of glaucoma. And so this has been a major focus of the Academy.

We are doing well in the -- if you look at the last 15 years, we've decreased51 blindness by 50 percent. The rate of blindness.

However, there is still marked differences in African Americans and Hispanics.

MEMBER GOLDBERG: Is the goal of this to get patients referred to ophthalmologists by their primary care doctors?

Yes, is the goal of this to get primary care physicians to refer to ophthalmologists? Or is this a measurement of -- for ophthalmologists?

DR. LUM: For ophthalmologists and optometrists. But it's really to improve the care and make sure that we are focusing on interocular pressure reduction.

Because we have shown studies and that's in the rationale of the Measure that you know, maybe 50 percent of patients aren't having their IOP reduced by 15 percent.

I think Dr. Stein has also had written articles to that effect. That when we look across and we look at patient records, we aren't seeing the interocular pressure reduction.

And that's the only -- as Vaishali said, the only modifiable risk factor that we can do anything about to affect the ultimate outcome of glaucoma and prevent blindness.

CO-CHAIR YAREMCHUK: Are we ready to vote?

MEMBER PATEL: I would like to add one more thing. To answer the question of, you know, is there a gap in care that warrants a National Performance Measure.

So, you know, even though we see high rates of reporting or meeting this Measure, you know, being 90 plus percent. Keep in mind that that -- about 70 percent of that is actual, you

know, meeting the 15 percent lowering of target.

Which I think is already very good. 70 percent, if people are meeting that criteria alone, 70 percent, that's already very good. But it can always be better.

So, that's why it's a good idea to continue to, you know, measure that. But, you know, don't -- I do think that the good portion of that 30 per -- you know, 95 percent is because of the second part of that wording which is Plan -- Document of Plan of Care.

So, if there is any way to separate the two in the future, you know, I would be in favor of that so that we can clearly see how many physicians are actually meeting the level of lowering the IOP versus how many times, you know, is a documented Plan of Care. And measure those two separately.

DR. LUM: And I think Dr. Rich had mentioned that we looked at our IRIS Registry data. This is the manual registry.

And 30 percent -- so for everybody who

claimed that they met the performance, 30 percent met it with plan of care. But 70 percent met it by reducing IOP.

So that probably reflects just that complicated situation, you know, that the patients come in, they're advanced age. Are you going to add another medication or surgery? Probably not.

You might have a plan of care of how to address it. The patient can't afford medication. You might have a plan of care.

You can't always add another medication or surgery to get that interocular pressure down in a year's time. And so, that's why that plan of care was provided.

But it still keeps their focus I think on reducing the interocular pressure. And looking out for that patient and monitoring that patient.

And that was the intent of the measure.

MEMBER PATEL: Yes. No, no, I -- so

I agree completely. I think that's a good idea. And I'm glad that in the IRIS Registry there is a way to separate the two.

I'm saying also, if in the quality measure, there is a way to separate the two, you know, for future purposes, that's something that should be considered.

DR. LUM: Well, certainly that can be separate. I don't -- CMS doesn't report it out. But there's different codes for reporting either a plan of care or a reduction of IOP.

So you can definitely separate it out.

CO-CHAIR YAREMCHUK: Are we ready to vote for opportunities for improvement?

MS. ROBINSON-ECTOR: Voting is now open for performance gap for Measure 0563. And for those on the phone, option one is high, two is moderate, three is low and four is insufficient.

All the votes are in. 64 percent voted high. 36 percent voted moderate. Zero voted low and zero voted insufficient. So for

performance gap, Measure 0563 passes.

CO-CHAIR YAREMCHUK: Okay. Reliability?

MEMBER STEIN: So I think for the most part, it's a reliable measure. You know, I think Vaishali brought up the issues of the plan of care.

And I think in a future iteration it would be nice to tease that out. I think that the devel -- and I don't want to speak for the developers, but I think their intention is to come up with a measure that where the goal is to get 100 percent.

And there may be ways to do the measure where the goal doesn't have to be 100 percent. And you can capture some of these other circumstances and have it where, you know, 90 percent should be, you know, should be good performance.

And maybe those who are getting 100 percent are over-utilizing or doing too much. But, I think for the most part it's a pretty

reliable measure.

You know, I think that I would encourage in the next iteration, for risk adjustment, clearly, with the increasing number of minorities with glaucoma, some of these patients are more difficult to care for. And achieving of 15 percent pressure lowering can be more challenging.

And you'd hate to have a provider who is practicing in a setting where you've got much more challenging patients not do as well because of their patient population and their demographics.

So, I think risk adjustment and hopefully with IRIS and -- will enable that to be possible.

DR. RICH: Yes. One of the issues of risk adjusting is it takes away the focus of the adverse outcomes of those populations at risk. So, you've got to balance risk adjustment with keeping the focus on the outcomes from people with -- Hispanics and African Americans.

But, we will be able to do that. We have eight million people in the Registry now. We're planning that probably 48 million by 2017. So we'll be able to look at large stratified data on what's unique about these populations with the plan of care.

And it -- that's actually not that hard to collect.

CO-CHAIR YAREMCHUK: Any other comments?

(No response)

CO-CHAIR YAREMCHUK: Are we ready to vote on reliability?

MS. ROBINSON-ECTOR: Voting for reliability for Measure 0563 is now open. And for those on the phone, option one is high, two is moderate, three is low and four is insufficient.

Okay. All the votes are in and voting is now closed. 43 percent voted high. 57 percent voted moderate. Zero voted low and zero voted insufficient. So, for reliability, Measure

0563 passes.

CO-CHAIR YAREMCHUK: Okay. Validity?

MEMBER STEIN: So for validity, the Developer used an expert panel, the 16 glaucoma specialists. And 15 of the 16 members, based on face validity, found it to be valid.

CO-CHAIR YAREMCHUK: Any other comments?

MEMBER MADONNA: Yes, I'd like to ask the Developers, although lowering IOP is obviously an admirable goal. Why was it 15 percent chosen as opposed to 25 percent? Or something else?

Even in the ocular hypertension treatment study, patients who have not developed glaucoma, there was a 20 percent goal IOP reduction. So, why 15?

DR. RICH: I'll let Flora start and I'll comment.

DR. LUM: So, you're right. And as Dr. Stein mentioned, the randomized control trials did show higher rates. But those are all

randomized control trials in a very particular setting.

You know, that may not translate into real world. As Dr. Stein's mentioned, there are minority populations. Populations that we're concerned about medication adherence. And that really affects the amount of interocular pressure reduction that you can have.

So it was really designed as kind of a floor -- a floor effect. A failure effect that you really -- 15 percent was -- really could be reasonably expected for the patient and on the part of the physician without going to extraordinary needs and having the most ideal patient that came in for every follow up visit.

MEMBER MADONNA: Well, I would think that the plan of care covers that. And if I were looking at my own patients and looked at 15 percent, I would think that they would be under-treated in almost all cases with the current state of the art with prostaglandin and analogues and so forth.

DR. RICH: Yes. The fact of the matter is that somewhere between 50 and 60 percent are not under control currently. So, you have to start someplace with the pressure lowering.

And again, you don't have -- it took cardiology eight years to develop the guidelines for hypertension working with other primary care groups. You know, we're trying to address, you know, a pressing need.

We can't wait eight years to develop new evidence based guidelines. Now they've got competing guidelines, but I won't get into that.

So, we thought that this was a reasonable Measure. If you look at other Measures, other intermediate outcome Measures in chronic disease, it's -- if you take a 15 percent cut of the vast majority of people with hypertension that are outside, you're going to have them inside your 140 over 90 guideline.

So, is there Level One evidence that this is the number? The answer is no. But, we

have a huge gap in care when somewhere between 50 to 60 percent of people under treatment have inadequate levels.

DR. STEIN: Richard, I think your point is well taken. I think one of the challenges is with caring for patients with glaucoma in general is that, you know, there's patient related factors.

There's provider related factors. And there's system related factors that are all going to impact what someone's pressure is on a given day.

And patient adherence is known to be a huge issue in glaucoma. So, how well one -- I think there needs to be some flexibility to deal with factors that are beyond the control of the provider.

I think that was part of the intention with choosing the 15 percent number.

DR. RICH: It resulted from extensive discussions with outcomes researchers like Josh and others in the glaucoma community. So, this

was a consensus figure.

But again, it's not an evidence based guideline like NHLBI originally came up with. You've got to start somewhere.

CO-CHAIR YAREMCHUK: Are we ready to vote? Or any more comments?

(No response)

CO-CHAIR YAREMCHUK: Okay. Ready to vote on validity.

MS. ROBINSON-ECTOR: Voting is now open for validity for Measure 0563. And for those on the call, option one is high, option two is moderate, three is low and four is insufficient.

All the votes are in. 7 percent voted high. 79 percent voted moderate. 14 percent voted low. And zero voted insufficient. So for validity for Measure 0563, the Measure passes.

CO-CHAIR YAREMCHUK: Feasibility?

MEMBER STEIN: I think that, you know, eye pressures are captured pretty well in electronic health records. So it's a reasonably

feasible, I mean Measure in my opinion.

CO-CHAIR YAREMCHUK: Any other comments?

MEMBER CARNAHAN: Yes, there's specific codes both for the 15 percent reduction as well as for the plan and care piece of it. So, it can be clearly documented.

And I think that's the good thing. I guess the other part of it, which is a different section, is, you can almost always find one of these plan of care options that would work should you not have the target pressure.

CO-CHAIR YAREMCHUK: Okay. Was there another comment on this side?

(No response)

CO-CHAIR YAREMCHUK: Okay. All right. Good. We're ready to vote on feasibility.

MS. ROBINSON-ECTOR: Voting on feasibility for Measure 0563 is now open. And for those on the phone, option one is high, option two is moderate, option three is low and four is insufficient.

All the votes are in. 64 percent voted high. 36 percent voted moderate. Zero voted low and zero voted insufficient. So for feasibility, Measure 0563 passes.

CO-CHAIR YAREMCHUK: Usability and use?

MEMBER STEIN: This data is captured in the IRIS Registry and electronic health records. And you know, with claims data it doesn't capture eye pressures.

But, it looks like we're moving away from using claims data as the other capture means. And it should be pretty usable.

CO-CHAIR YAREMCHUK: Any other comments?

DR. WINKLER: I have one question. Is there any plan to publically report these Measures? Since these are from your Registry and not -- this one's not part of the federal programs.

DR. LUM: Everything in PQRS will be reported, right? In 2015 for -- yes.

DR. RICH: Next year. In 2015.

DR. LUM: Yes.

CO-CHAIR YAREMCHUK: Okay. Ready to vote.

MS. ROBINSON-ECTOR: Voting is now open for usability and use for Measure 0563. And for those on the call, option one is high, two is moderate, three is low and four is insufficient information.

Okay. All the votes are in. 43 percent voted high. 50 percent voted moderate. 7 percent voted low. And zero voted insufficient. So for usability and use, Measure 0563 passes.

CO-CHAIR YAREMCHUK: Okay. Now, vote on whether to recommend that Measure as suitable for endorsement.

MS. ROBINSON-ECTOR: Voting is now open for recommendation for overall suitability for endorsement for Measure 0563. For those on the call, option one is yes and option two is not.

And it looks like we're missing one vote in the room. If you all could revote please?

Okay. Now all the votes are in. 93 percent voted yes. 7 percent voted no. So for recommendation for overall suitability for endorsement for Measure 0563, the Measure passes.

CO-CHAIR YAREMCHUK: Okay. Thank you. And this isn't an E-Measure. So we don't need to vote separately for it.

The next Measure is Primary Open Angle Glaucoma, 0086, Optic Nerve Evaluation. Evidence? Pardon me? Okay. Sheli?

MS. LUONG: Vaishali, would you like to speak to this Measure?

MEMBER PATEL: The optic nerve damage Measure?

MS. LUONG: Yes. As a -- the 0086. The optic nerve evaluation.

MEMBER PATEL: yes.

MS. LUONG: Yes.

MEMBER PATEL: So, --

MS. LUONG: Just the evidence for now.

MEMBER PATEL: Yes. Give me one minute. So, regarding the evidence to support the Measure, the evidence, it cites the IOP practice pattern as assessing guideline -- you know, evidence Grade A for the Measure, which I agree with.

And then also, there is evidence that's cited in the IOP practice pattern that's based on -- it's actually, you know, based more on the case series rather than the large randomized controlled trials.

But there are enough large randomized controlled trials that also use the optic nerve damage as an assessment for glaucoma. So I would say the evidence is there to support the Measure.

CO-CHAIR YAREMCHUK: Any comments?

DR. RICH: One other -- I'll just add a quick note. And that is, when we were looking to address, there is a disparity.

About ten years ago we tried to find a "cheap, less invasive way" of looking at

populations at risk. This was -- and so, in detailed by an examination of the optic nerve was shown by Hopkins -- by Dunbar Hoskins in the '90s to actually be able to identify people at risk.

So, there was a glaucoma detection benefit was passed with the -- for people -- African Americans and Hispanics in the family history of glaucoma to assess their possibility of having glaucoma.

And it was to preventative benefit designed by ourselves, the National Eye Institute, the American Glaucoma Society. The validity was approved by CMS and the CBO. And the cost savings were scored profit -- positively by CBO.

So, low tech is good.

CO-CHAIR YAREMCHUK: So, are we ready to vote evidence?

MS. ROBINSON-ECTOR: So, voting is now open for evidence for Measure 0086. And for those on the phone, option one is high, two is moderate, three is low and four is insufficient

evidence.

Okay. All the votes are in 43 percent voted high. 50 percent voted moderate. 7 percent voted low. And zero voted insufficient evidence. So for evidence, Measure 0086 passes.

CO-CHAIR YAREMCHUK: Okay. Opportunity for improvement?

MEMBER PATEL: So, the current reporting for the Measure is in the 90 plus percent range. But, I think it's contin -- you know, there is still opportunity for improvement toward 100 percent.

And I also think, as it was mentioned before in the other, you know, Measure, it's important to continue to measure this so that it remains high.

DR. RICH: Another comment. There's previous literature showing the other -- the failure to document the optic nerve and this is a chronic disease.

It would be nice in hypertension to look at what's happened in the blood vessels in

the heart. Well, we can do that looking in the eye.

And second of all, the performance rate was really only 79 percent when we actually looked at charts and looked at the electronic record. Flora?

DR. LUM: Yes, that percentage was much lower when we actually looked at it. And we've talked to practices and they don't.

The two things that we look for in the E-Measure specification, which are also in the Measure specification for registry, is cup to disc ratio and documentation of optic nerve appearance. And we talk to practices all the time and they don't document that.

So, that's -- that is a gap for improvement.

MEMBER PATEL: Okay. So you're saying even though it's reporting PQRS as 90 plus percent, when you looked yourselves you found it be lower?

DR. LUM: Right. Because as you know,

that's attestation. So if they put a claims code that said they did a dilated eye exam, which what the bare minimum of the Measure specification says.

But we actually in the Measure rationale we say when you look at the nerve, you need to document -- look at the cup to disc ratio and document the optic nerve appearance.

And those will really give you a guide on how to treat the patient. Whether management needs to be added or, you know, stepped up.

So, those are the things that we look for in the electronic health record. And we don't find them really, well, a higher percentage of the time.

CO-CHAIR YAREMCHUK: Todd?

MEMBER RAMBASEK: So, if I'm understanding correctly, you can do an optic nerve head evaluation and not document the cup to disc ratio?

DR. LUM: Well, yes. That's what we're saying. Is in the claims and registry

process, the doctors are saying that they've done it. They probably in, you know, they think they've done it.

But if you actually look at those who have electronic health records, there is no documentation of the cup to disc ratio or appearance.

MEMBER RAMBASEK: So, is that what the Measure should be? That if you have a diagnosis of POAG that you need to document the cup to disc ratio?

DR. LUM: It is in the Measure, in the details of the Measure description from which the E-Measure specification has it. We just are very -- were much more precise in what we extract from the EHR.

MEMBER FRIEDMAN: So, just a quick point of clarification. So if they say they doc -- they say they looked at it, how did they document that they looked at it?

I mean, you can write down, I looked at the optic nerve. But is there any evidence

that they actually did do what they said they did?

DR. RICH: Well, it's pretty good -- this was one that was not registry reported. It could be reported on claims. So it was an attestation.

Our point is, when you actually look at the performance electronically, it's not there. So, there's a huge area for improvement and a gap in care.

So, you know, no one's going to be reporting this on claims this year. They're only going to be reporting it electronically.

And so, the performance is overstated. But what we've learned in the claims for this particular Measure. And there are some problems that we'll get into about the Electronic Measure later on.

MEMBER MADONNA: Yes, I was just going to ask, the Measure says optic nerve head evaluation. Are they evaluating it by fundus photography and making a comment about the fundus

photos? Or by OCT?

MEMBER PATEL: Am I on? I cannot hear whatever is being discussed with the phone.

CO-CHAIR YAREMCHUK: No one's saying anything right now. We're waiting for someone to come up to the microphone.

MEMBER MADONNA: Differently, it is appropriate to get an OCT which has a cup to disc ratio on it and utilize that instead of actually -- because this just says optic nerve head evaluation.

And that could be done a number of different ways.

DR. LUM: So, as mentioned before, it includes the cup to disc ratio and the structural -- the exam for structural abnormalities.

DR. WINKLER: But in the code set for the numerator like what for an examination, don't the optic nerve photographs and OCT codes also count in that as a way to document the optic nerve?

DR. LUM: I think we do -- I mean,

that is part of the preferred practice pattern, is that we do tell the ophthalmologist to look at the optic nerve by dilated eye examination. In addition, they can do ancillary testing.

It's considered part of the eye exam, yes.

MEMBER MADONNA: Yes. Even though I'm kind of almost contradicting what I said, I mean, I do think it's important that you look -- that you actually look at the optic nerve head because of the presence or absence of trans-hemorrhages.

So, in that sense, I really do think you should. But there seems to be -- and it seems to be in this wording that optic nerve head evaluation could be done a number of different ways.

MEMBER STEIN: You know, we've looked at claims data. And I think this 95 percent for in terms of performance is a way over estimate.

In claims data it's more in the 30 percent range. And I think part of it is, is the devil's in the details, as you were getting to,

as to what constitutes an examination of the optic nerve head?

You know, does doing an OCT test count? Does -- how descriptive a definition is it? Is it just a check off sheet, yes, I did it?

So, I think it's important to clarify exactly what counts and what does not count.

MEMBER GOLDBERG: Do you include 92133 and 92134 in your data collection?

DR. RICH: Did you throw those number out as my brain is befuddled?

DR. GOLDBERG: 92133 is for ophthalmic diagnostic imaging post resegment of the optic nerve. And then 92134 includes the retina.

Are you including that in your --

MEMBER STEIN: It's the 331 that's relevant here. The 341 is for diseases of the macula.

CO-CHAIR YAREMCHUK: Any other comments? Questions?

(No response)

CO-CHAIR YAREMCHUK: And are we ready

to vote?

MS. ROBINSON-ECTOR: Voting is now open for performance gap for Measure 0086. And for those on the call, option one is high, two is moderate, three is low and four is insufficient.

All the votes are in. 50 percent voted high. 50 percent voted moderate. Zero voted low and zero voted insufficient. So for performance gap for Measure 0086, the measure passes.

CO-CHAIR YAREMCHUK: Reliability?

MEMBER PATEL: Okay, sorry, I was -- I'm here. Sorry, I was on mute.

So, I would say we have all of the stated.

MS. LUONG: Vaishali, just to know, right now we're talking about reliability. What the claims and registry version of Measure 0086.

MEMBER PATEL: Okay.

MS. LUONG: And then we'll talk about reliability and validity together for the E-Measure version later. Thanks.

MEMBER PATEL: Got it. Okay. Yes. So, I would say there was extensive reliability information provided by the Developer, which is great to see.

Reliability was measured using, you know, claims, the PQRS data set as well as the registry. And reliability seems high from the claims data even though, you know, it's important to note that not all -- there were a total of 45 thousand physicians in the data set.

Of which only nine, you know, about ten thousand were eligible for the reliability testing. So, that's I would say is a limitation. But I can see the reason for doing that.

And likewise from the registry, it's -- there's also reliability information provided. So I would say the reliability data is high quality.

CO-CHAIR YAREMCHUK: Any questions? Comments?

(No response)

CO-CHAIR YAREMCHUK: Are we ready to vote on reliability?

MS. ROBINSON-ECTOR: Voting is now open for reliability for the claims and registry version of Measure 0086. And for those on the call, option one is high, two is moderate, three is low and four is insufficient.

Okay. All the votes are in. 43 percent voted high. 57 percent voted moderate. Zero voted low and zero voted insufficient. So for the claims and registry version of Measure 0086 for reliability, the Measure passes.

CO-CHAIR YAREMCHUK: Validity?

MEMBER PATEL: For the face validity of the Measure was assessed by an expert panel of 16 members. And -- who agreed that this Measure could distinguish quality of care.

And, I would agree with that approach for assessing face validity. Of course, I'll repeat that it's nicer to see a larger sample. But, I think that this is good.

CO-CHAIR YAREMCHUK: Any comments?

(No response.)

CO-CHAIR YAREMCHUK: Ready to vote on validity?

MS. ROBINSON-ECTOR: Voting for validity for the claim and registry version of Measure 0086 is now open. For those on the call, option one is high, two is moderate, three is low and four is insufficient.

All the votes are in. 21 percent voted high. 79 percent voted moderate. Zero voted low and zero voted insufficient. So for validity for the claims and registry version of Measure 0086, the measure passes.

CO-CHAIR YAREMCHUK: All right. Moving on. So, next is to be discussed is feasibility.

MEMBER PATEL: So I would say this data can be captured as we discussed, from claims and from registry. So it's feasible to capture.

And I would say we've already discussed the limitations of capturing it from the claims. Regarding the legal of granularity

we can have with the data, it would actually be better to capture this from registry then claims is the only comment I would make.

But it's feasible.

CO-CHAIR YAREMCHUK: Comments?

(No response.)

CO-CHAIR YAREMCHUK: And I guess I just have a question. Are we capturing the right data? Because I've heard it go back about cup to disc ratio.

We've talked about other I'm going to say descriptive terminology versus just that they did the exam. And the question is, in this Measure, are we going to be capturing the data that we feel comfortable with?

DR. LUM: So as we discussed, the Measure rationale does talk about what needs to be looked at in the dilated eye evaluation. And refers back to the preferred practice patterns which talks about the characteristics of the optic nerve and what should be noted.

And that's the appearance and the cup

to disc ratio are the two salient things. So that's already in the Measure rationale.

And hopefully when people do report it, you know, that they are doing those things.

DR. RICH: Flora, can you just --

MEMBER PATEL: Yes, and you know, I would comment on that also. To say that, you know, even if we can get patients to come in for the required number of times per year and we can even document that an optic nerve head exam was done that's already a win.

Because a lot of times the issue may be that patients are not even coming in.

DR. RICH: Just to further answer your question, Rich raised the thing, does a photo, and the answer is no. We want you to look. And then if you want to follow them longitudinally.

So, the codes for the OCT are not included. So, we are capturing what we thought. But we just had -- it's late in the day and we had to look it up.

CO-CHAIR YAREMCHUK: Josh?

MEMBER STEIN: Do you think that's good thing?

DR. RICH: Well, sometimes it is. Sorry. The Hoskins paper suggests that that's adequate. Now, for longitudinal follow up, I think that, you know, probably photos and OCT are appropriate.

MEMBER STEIN: And then one other comment. I think getting at a point raised earlier, you know, getting patients in, since there are differences in access to care among patients of different sociodemographic profiles, I would encourage the next iteration for risk adjustment.

Because clearly, we're seeing this in the claims data, patients of certain profiles don't follow up for regular exams as often as others. And you don't want the provider to be impacted based on, you know, the types of patients they're caring for.

DR. LUM: Okay. I was just going to add that the dilated eye evaluation and

documentation, we know that OCTs and photos are useful for follow up. But this is only required once over the whole year.

So we're saying that, you know, there should be a look into the eye at least once a year for that patient. No matter the other follow ups and ancillary testing.

DR. RICH: And as you pointed out, you and Paul in previous papers, there's a huge gap in care that the optic nerve is looked at. I think our number of 79 percent is surprising to us.

CO-CHAIR YAREMCHUK: Go ahead.

MEMBER MADONNA: You mentioned that the preferred practice patterns mention cup to disc ratio and a description of the nerve. But what's actually captured when we capture the data?

Is it just the cup to disc ratio because that's the easiest thing to pull of the EHR?

DR. WINKLER: Right now we're talking

about the claims and registry version. Not the EHR. So, what's the data element specifically for the claims and registry?

DR. LUM: Okay. It's the dilated look at the optic nerve. Right. It's a G Code that the --

MEMBER MADONNA: Just whether that it was done or not.

DR. LUM: Right.

CO-CHAIR YAREMCHUK: Okay. Are we ready to vote on feasibility?

MS. ROBINSON-ECTOR: So, the voting is open for feasibility for the claims and registry version of Measure 0086. For those on the phone, option one is high, two is moderate, three is low and four is insufficient.

All the votes are in. 21 percent voted high. 57 percent voted moderate. 21 percent voted low. And zero voted insufficient. So for feasibility of the claims and registry version of Measure 0086, the Measure passes.

CO-CHAIR YAREMCHUK: Okay. Usability

and use?

MEMBER PATEL: So, the Measure is obviously currently in use. And is shown to be usable, you know, in the PQRS program. And also the EHR -- well, I guess I won't say with the registry.

But the PQRS program and also reported in the IRIS registry.

CO-CHAIR YAREMCHUK: Any comments from anyone?

MEMBER STRODE: At risk of being repetitious, but for my clarity, the aim is to see that at least there's a description once a year of a cup to disc ratio and the appearance of the disc.

But what's actually being captured in the registry is simply the attestation. I looked at the back of the eye or I didn't.

DR. LUM: That is correct for the registry and claims. Obviously if you were to be -- if the physician were to be audited, they would have to provide evidence in the medical

record that they did look at the optic nerve.

And they, you know, we would expect that there would be documentation of its appearance. I think that's what the auditors would look at if they were looking if they were to audit the records.

DR. RICH: And in the registry, we actually have a chart to do data validity and auditing. So we -- and we do have the ability to actually see the physician's electronic record and what was recorded.

MEMBER MADONNA: Does this in some way negatively impact the performance of visual field testing? Particularly in patients who have advanced disease where looking at the nerve is probably is little utility, where a field has much more utility.

And someone sitting very close to me, I know, published the paper about the reduction in the use of visual field testing, correct?

Yes, so. Anyway.

DR. RICH: I'll defer to Josh. But --

MEMBER STEIN: Well I think it -- as a glaucoma specialist, I think it's important both to assess the optic nerve and assess a visual field. And maybe there's a measure that's under development to look at rates of visual field testing.

I don't think we should knock this Measure because it doesn't include visual field. But, I think both are important. And there are probably a small subset of patients that either you're not going to get that much additional information either from assessing the nerve or doing the field.

Because some patients you just can't get a reliable field on. In some patients, your view of the optic nerve may not be that good.

But I think designing a Measure that's going to capture the majority of patients that mo -- I think most of us would agree that the majority of glaucoma patients, we should be looking at their nerve.

CO-CHAIR YAREMCHUK: All right. We're

ready to vote for usability and use.

MEMBER PATEL: Yes.

MS. ROBINSON-ECTOR: Voting is now open for usability and use of the claims and registry version of Measure 0086. For those on the call, option one is high, two is moderate, three is low and four is insufficient information.

All the votes are in. 57 percent voted high. 36 percent voted moderate. 7 percent voted low. And zero voted insufficient. So for usability and use of the claims and registry version of 0086, the Measure passes.

CO-CHAIR YAREMCHUK: Okay. Next is for the Committee to vote on whether to recommend Measure for endorsement.

MS. ROBINSON-ECTOR: Voting is now open for recommendation for overall suitability for endorsement for the claims and registry version of 0086. For those on the call, option one is yes and option two is no.

Okay. All the votes are in. 93

percent voted yes. And 7 percent voted no. So for the recommendation of overall suitability for endorsement for the claims and registry version of Measure 0086, the Measure passes.

DR. WINKLER: Go to the next slide. All right, this Measure also does have an E-Measure version. And I think if we think about the scientific acceptability, the issues around specifications.

So, I just want to go back to the comments you've made already about how you compared the claims to the information in the EHR. And so, the EHR specification -- E-Measure specifications are more detailed or more specific then the claims and registry.

Could you just describe that a little bit more clearly? So we understand that.

DR. LUM: That's correct. So, in the EHR specification, we have to look for cup to disc ratio specifically in the record. And we have to look for an optic nerve appearance that's documented in the record.

So that differs, you know, as we said, in terms of the registry.

DR. RICH: Yes. And that's going to be the operative way of collecting this Measure. Because it's not going to be done by claims this year.

MEMBER STEIN: Flora, can you clarify? So, if a provider is drawing a picture of the nerve in the EHR, that is number one capturable? And number two, will get credit?

DR. LUM: Yes, that's right.

MEMBER STEIN: Okay.

DR. WINKLER: Though similar to the other E-Measures in the past, we combined reliability and validity into scientific acceptability with the same thing. We're going to see the results from testing on the simulated data set.

MS. ROBINSON-ECTOR: So, voting is now open for reliability and validity for the E-Measure version of 0086. And for those on the call, option one is yes with conditions and

option two is no.

Okay. All the votes are in. 100 percent voted yes with conditions. And zero percent voted no. So for reliability and validity for the E-Measure version of 0086, the Measure passes.

DR. WINKLER: All right. So we'll also talk about the feasibility of the E-Measure. Any questions or issues about that?

(No response)

DR. WINKLER: So we'll vote on that one.

MS. ROBINSON-ECTOR: Voting is now open for feasibility for the E-version of 0086. And for those on the call, option one is high, two is moderate, three is low and four is insufficient.

Okay. All the votes are in. 64 percent voted high. 36 percent voted moderate. Zero voted low and zero voted insufficient. So for feasibility of the E-Measure version of 0086, the Measure passes.

DR. WINKLER: Okay. So for usability and use, there were some differences that we just talked about between the claims and registry.

So, perhaps it would be good just to take the vote on usability and use specifically for the E-Measure so that you could potentially vote differently, I think.

MS. ROBINSON-ECTOR: So, voting is now open for usability and use for the E-version of 0086. And for those on the call, option one is high, two is moderate, three is low and four is insufficient.

All the votes are in. 36 percent voted high. 64 percent voted moderate. Zero voted low and zero voted insufficient information. So for the E-version of 0086 for usability and use, the Measure passes.

DR. WINKLER: Go ahead.

MS. ROBINSON-ECTOR: Okay. Voting is now open for recommendation for overall suitability for endorsement for the E-Measure version of 0086. And for those on the call,

option one is yes with conditions and option two is no.

Okay. All the votes are in. So 100 percent voted yes with conditions. And zero percent voted no. So, for recommendation for overall suitability for endorsement for the E-Measure version of 0086, the Measure passes.

MS. LUONG: Great. So we ended our review of the Measures for today on -- five minutes early.

Operator, can you open the line up for our member and public comments on the phone? Or if anyone wants to speak in the conference room, please feel free to do so.

OPERATOR: All right. At this time, if you would like to make a public comment, please press star then the number one on your telephone key pad.

And there are no public comments at this time.

MS. LUONG: Well, we just wanted to thank everyone on behalf of NQF for being so

courteous with the timing and all the technical difficulties. And for our thoughtful discussions for day one.

We do have reservations in place for a restaurant, which I'll email everyone for 6:00 p.m. today for the Standing Committee Members. It will be a good chance for you all to get together to interact and to talk about your day.

So, I'll email out the location now. If you guys can make it. Can I just get a show of hands just to see?

Who -- it's like two blocks away. Happy hour and dinner.

Yes, for both.

CO-CHAIR YAREMCHUK: I think you got a few more votes.

MS. LUONG: Yes. Sounds good. It's two blocks away. Right down the block.

So, I'll send out that information now. Thank you everyone.

MEMBER PATEL: Thank you, everyone.

DR. WINKLER: All right. Tomorrow

morning we'll start just a tad earlier. If you notice on the agenda, just so that everybody gets here on time.

The continental breakfast is at 8:00 and we'll reconvene for discussion at 8:30. Okay. So it's a half an hour earlier.

Thank you all very much. Have a good evening. We'll see you tomorrow.

MS. LUONG: And thank you, the Developers as well.

(Whereupon, the above-entitled matter went off the record at 4:42 p.m.)

H

N

W

C E R T I F I C A T E

This is to certify that the foregoing transcript

In the matter of: Eye Care, Ear, Nose and Throat Conditions (EENT) Standing Committee

Before: NQF

Date: 06-03-2015

Place: Washington, D.C.

was duly recorded and accurately transcribed under my direction; further, that said transcript is a true and accurate record of the proceedings.

Neal R Gross

Court Reporter