National Quality Forum
Primary Care and Chronic Illness Steering Committee
Thursday, June 23, 2022

The Committee met via Videoconference, at 9:00 a.m. EDT, Dale Bratzler and Adam Thompson, Co-Chairs, presiding.

Present:

Dale Bratzler, DO, MPH, Co-Chair
Adam Thompson, BA, Co-Chair
Ann Kearns, MD, PhD, Mayo Clinic
Anna McCollister, Galileo Analytics
Carlos Bagley, MD FAANS, UT Southwestern
Grace Lee, MD, Virginia Mason Medical Center
James Mitchell Harris, PhD, Children's Hospital Association
James Rosenzweig, MD, Hebrew SeniorLife
Kim Elliott, PhD, Health Services Advisory Group
Lindsay Botsford, MD, Memorial Hermann Medical Group
Robert Bailey, MD, Johnson & Johnson
Starlin Haydon-Greatting, MS-MPH, BSPharm, CDM, FAPhA, SHG Clinical Consulting
William Curry, MD, Penn State College of Medicine
William Glomb, MD, FCCP, FAAP, Superior HealthPlan

Surgery Standing Committee Members Present:

Richard D'Agostino, MD
Michael Firstenberg, MD, FACC, FAIM
Vilma Joseph, MD, MPH, FASA
Miklos Kertai, MD, PhD
Salvatore Scali, MD, FACS, DFSVS, RPVI

NQF Staff:

Leeann White, MS, BSN, Director
Taroon Amin, PhD, Consultant
Matilda Epstein, MPH, Associate
Matt Pickering, PharmD, Senior Director
Victoria Quinones, AA, PMP, Project Manager
Isaac Sakyi, MSGH, Manager
Tristan Wind, BS, ACHE-SA, Analyst

Also Present:

Naomi Bardach, MD, University of California,

San Francisco
Michael Cabana, MD, MPH, The Center for Health and Community
Collette Cole, MN Community Measurement
Michael Firstenberg, MD, Ascension NE Wisconsin-St. Elizabeth
Mitch Harris, Harris Tech Strategies
Jeff Jacobs, MD, FACS, FACC, Congenital Heart Center
Julie McCormick, University of Michigan
Dave Shahian, MD, Mass General Research Institute
Julie Sonier, MN Community Measurement
Moritz Wyler von Ballmoos, MD, Houston Methodist
Banu Yagci, The Society of Thoracic Surgeons

Contents

Proceedings

(9:02 a.m.)

Welcome and Review of Meeting Objectives

Ms. White: Okay, well, it is 9:02 a.m. on the east coast. So good morning, everyone. Welcome. Greetings.

My name is LeeAnn White and I'm the director supporting the Primary Care and Chronic Illness Project for the Spring '22 cycle.

So I'm glad to see everyone is here. Happy to meet most of you for the first time. This is a newer team for this cycle so I'm really excited to work with you all this spring.

I first want to also thank you for your time and participation. I do understand that it's a significant amount of time and effort to review these measures and prepare for today's review.

I'd also like to extend a thank you to our developers for being on the call today. We recognize that there is significant time and effort that goes into the testing, the preparation of materials, and the measure submissions.

So we definitely want to highlight those efforts and thank them for their time as well.

And then lastly, I'd like to also share my appreciation for your continued patience and understanding as we continue to meet virtually in the pandemic.

We do understand those challenges that accompany virtual meetings, and we all look forward to the time when we can meet in person.

However, in the meantime, we do appreciate your understanding and thank you for your continued support.

So we're going to try to make this as -- bridge that gap as much as possible and try to make this more of an in-person deal across those miles.

So I'm going to have -- I'm going to open it up and let our esteemed co-chairs, we have Dr. Dale Bratzler and Adam Thompson, and I'd like to offer them the opportunity to provide their welcoming remarks. So, Dale, Adam?

Co-Chair Bratzler: Well, I'll start, just thank you everybody for participation. I was telling earlier that I had a chance to read through all of the work that you did on the measures beforehand.

I really appreciated all the comments that you put into the document. So we should have a very robust conversation today. And I appreciate the work you're doing.

Co-Chair Thompson: I echo that. Thank you for all the work and welcome to our colleagues from other committees. It's great to have you here as part of our committee family today.

Also, I just want to let folks know, I will have to drop off around 2 o'clock PM today. So my apologies that I'll miss the end of our meeting, but you will be in Dale's very competent hands. So thank you all.

Ms. White: All right. Wonderful. Thank you so much. Okay, so we're going to -- with the virtual meetings, we're going to take a pause here to get our slides up on our screen.

So if you'll just bear with us just a moment, we will get our slide deck up.

Just a reminder, a couple of reminders. So the meeting invite contains all the materials, the slide deck and the agenda for today.

All right. Perfect. Thank you so much. So we will go to the next slide, please. And one more slide. And we'll go through housekeeping. Perfect. Thank you

so much.

So I want to take a brief moment to go quickly through a couple of housekeeping reminders.

As most of you know, we are using the Webex platform to host the Measure Evaluation meeting today.

If you're having any technical difficulties, please let our team know. You can go through our chat function with WebEx, so you can chat us directly or you can email us at our project team inbox. So it's primarycare@qualityform.org.

And in the spirit of engagement and collaboration, we also encourage you to use your video so that we can see each other's faces and bridge some of those virtual gaps.

If you're not actively speaking, we do ask that you please place yourself on mute to minimize any background noise and interruptions.

To mute and unmute, you just click on the microphone button at the bottom of your screen. We do highly encourage everyone to use the chat box feature and the raised hand feature throughout the meeting today.

NQF staff and the co-chairs will monitor the discussions and highlight comments throughout the call.

And there is also an option to speak or chat with people directly through the chat function.

So you can go to the whole group using the Everyone, or directly message someone through the dropdown function.

With the raise hand feature, the raise hand feature will alert the standing committee, the co-chairs, and NQF staff.

And so the raise hand icon will appear on your screen. To raise your hand, just click on the participants list and you'll find your name and you'll see that raise hand icon.

To unraise your hand, or to lower your hand, you click on that raise hand icon again.

If you're on the line and you've dialed in on your phone, to unmute and mute, you just do *6. So I just wanted to mention that for the phone participants.

One the meeting begins, our senior Director of Measurement Science and Application, Dr. Matt Pickering, will conduct roll calls and review disclosures of interest.

It is important to note that we are a voting body, and therefore we need to establish a quorum to vote on our meeting today.

If you need to step away from the call, we do ask that you please send our NQF staff a direct message using chat so that we are aware of the attendances and quorum. So next slide, please.

So it's my pleasure to go ahead and introduce our project team. Again, my name is LeeAnn White and I'm the director supporting this project.

Pictured here is our team manager, Isaac Sakyi, our analyst, Tristan Wind, our associate, Matilda Epstein, and then our project manager, Victoria Quinones.

We also have additional support staff present on the call to help address any of our questions and provide additional support.

Our senior director, Matt Pickering, and our project consultant, Dr. Taroon Amin. Next slide, please.

Some of the agenda items that we have listed here on our screen and that we'll be covering today, we

are going to begin again by conducting that roll call and disclosures of interest.

A Measured Specific Disclosure of Interest form was sent to each member as a standing committee.

We must receive this form to review any potential conflicts of interest. Unfortunately, if we have not received this form from you, you will not be able to participate in the discussions or voting today.

If you have an outstanding MSDOI, we will reach out to you directly and send you that form.

We do ask that you please fill out that form and send that back to us promptly so that you can participate in the call.

After we go through the disclosures of interest, Isaac will provide an overview of the evaluation process and the voting process. Then Tristan will conduct a quick voting test.

We did send out an email at approximately 8:30 Eastern time, this morning, that contains the voting link.

If you are unable to find that email, please let our staff know and we can definitely resend that.

We are not able to place it in chat. It's only for the Sandy community members. So we will work with you offline to get you that link so we can go through the voting chat.

We are using Poll Everywhere, which is our online platform for the voting. After the voting test, I'll briefly introduce our measures under review and then hand the discussions over to our co-chairs to facilitate our discussions.

Within that discussion, each criterion and voting on each criterion, we also then want to notify you that NQF has created a designated time frame for developers to respond to questions and provide

clarification.

The co-chairs and staff will collect any questions from the developer during the discussion for each criterion.

This does include those questions that are placed in the chat. And then the developers will be given the opportunity to respond to those questions and clarify any information prior to the standing committee vote.

The last vote we will have will be the overall recommendation for endorsement of the measure.

Following our measure discussion, we will review related NQP measures, and then we will host an opportunity for NQF members and public comments.

We will conclude with next steps and then we will adjourn our call. So next slide, please.

Okay, with that, I will hand this on over to Dr. Matt Pickering who will conduct roll call and review our disclosures of interest. So, Matt?

Dr. Pickering: Thanks, LeeAnn. Can you hear me okay? LeeAnn, can you hear me?

Ms. White: Yes.

Dr. Pickering: Good. Excellent. Great. Thank you. So we'll go to the next slide. So thank you everyone for your time, as LeeAnn said.

As we go through virtual meetings, especially these all-day types of meetings, we understand it's a lot to go through in a virtual environment.

So we do appreciate your time and always your attention to our work.

Today we'll combine introductions with disclosures of interest. So as LeeAnn mentioned, you received two disclosures of interest forms.

One is for our annual disclosures of interest that goes out every year to everyone and the other is really specific to the measures that we'll be reviewing this cycle.

So I know as far as we've asked a number of questions about your professional activities, and today we'll ask you to verbally disclose any of that information provided on those forms, and if you believe you have any potential conflicts, please state so as well.

We're especially interested in any grants, research and consulting related to this committee's work.

And just a few reminders. You sit on this group as an individual. You do not represent the interests of your employer or anyone who may have nominated you for this committee.

We are interested in your disclosures of any paid or unpaid activities that are relevant to the work in front of you.

And finally, just because you disclose does not mean you have a potential conflict of interest. We do verbal disclosures in the spirit of openness and transparency.

Now, we'll go around this virtual table, starting with our committee co-chairs, and I'll call your name.

So please then state your name, what organization you are with, and if you have anything to disclose.

If you do not have any disclosures, please just state that I have nothing to disclose to keep us moving along.

If you experience trouble unmuting yourself, please raise your hand so that our staff can assist you with that.

So as I go down the names, please, I apologize if I mispronounce your name in any way. I do apologize

about that. But again, we'll start at the very top. So Dale Bratzler?

Introductions and Disclosures of Interest

Co-Chair Bratzler: Hi, I'm Dale Bratzler, University of Oklahoma. I am here and I have nothing to disclose.

Dr. Pickering: Great. Thank you, Dale. Adam Thompson?

Co-Chair Thompson: Yes, Adam Thompson here with the Northeast Caribbean AIDS Education and Training Center.

And I do consulting work with USCF but have not worked with this particular measure in any way.

Dr. Pickering: Thank you, Adam. And then Ann Kearns?

Member Kearns: Yes, I'm here. Ann Kearns. I'm at the Mayo Clinic in Rochester.

Dr. Pickering: Thank you, Ann, and Kearns, apologies about that. Anna McCollister? Anna McCollister?

Okay. Moving to the next. Carlos Bagley? Carlos Bagley? Okay. Grace Lee?

Member Lee: Hi, I'm with the Virginia Mason Medical Center and I have nothing to disclose.

Dr. Pickering: All right, and Grace, I think we are potentially missing the MSDOI or the Measure Specific Disclosure of Interest form from you.

So the team will be sending it out to you. So please keep an eye out in your email. In order for us to have you participate in the voting, we'll have to have that submitted from you.

But you've verbally stated you have nothing to disclose, we just need to have that in the form. So

we'll be sending that to you.

Member Lee: Great.

Dr. Pickering: Thanks, Grace. And James Mitchell Harris?

Dr. Harris: Hi, Mitch Harris. I work at the Children's Hospital Association. A number of years ago, I did some work with QMetric at University of Michigan, who does have a measure up for a maintenance review today, but it's been over five years.

Dr. Pickering: Okay. Great. Thank you, Mitch. Appreciate that. And so being over that five-year limit, that wouldn't require any of that recusal.

So thank you. James Rosenzweig? My apologies about that. James, are you on? Okay. Kim Elliott?

Co-Chair Thompson: I think he's unable -- I think he's trying to get off mute.

Dr. Pickering: Oh, James, are you there?

Co-Chair Thompson: He's waiving at us. It looks like he's on mute.

Dr. Pickering: So, James, we'll circle back to you. See if we can try to get you off mute. Thanks, Adam, for pointing that out. We'll come back to James. Kim Elliott?

Member Elliott: Hi, Kim Elliott with Health Services Advisory Group, and I have nothing to disclose.

Dr. Pickering: Great. Thanks, Kim. Lindsay Botsford?

O: Good morning.

Dr. Pickering: Lindsay? Yes, go ahead.

Member Botsford: Yes. Lindsay Botsford, family physician in Chiefland with One Medical.

Dr. Pickering: And do you have anything to disclose, Lindsay?

Member Botsford: Nothing to disclose.

Dr. Pickering: Okay, thank you. Robert Bailey?

Member Bailey: Hi, good morning, Bob Bailey, Chapter of Scientific Affairs, and I'm a stockholder of Johnson & Johnson.

Dr. Pickering: Great. Thank you, Bob. Starlin Haydon-Greatting?

Member Haydon-Greatting: Excellent. Yes. It looks like Great-ing but it's pronounced Greatting.

I am a pharmacoepidemiologist, health economist. I have my own consulting practice, SHG Clinical Consulting.

I consult with the Illinois Pharmacists Association and other pharmacy associations. But what I probably need to disclose is for 23 years,

I was the Director of Quality Assurance for Medicaid, 50 percent Fed and 50 percent Illinois.

I'm still on retainer to be on their advisory panel. So only for their patient review. But I thought that since we're talking about California Medicaid data, I felt that you guys needed to know that I had 23 years of experience of looking at Medicaid data.

So, anyway, but that's the only thing that I have to disclose. Thank you.

Dr. Pickering: Thank you, Starlin, for disclosing that. Shouldn't be an issue with any work we're doing as far as any recusals for you.

So thank you for disclosing that. And good to talk with you again. It's been a while.

Member Haydon-Greatting: It has.

Dr. Pickering: Thank you very much. William Curry?

Member Curry: Hi, I'm Bill Curry. I'm a family physician at Penn State College and that is in Hershey, Pennsylvania.

The only thing I need to add is that recently we were awarded an NCI grant for lung cancer screening in primary care that will start in July.

Dr. Pickering: Okay, thank you very much for disclosing that. And then William Glomb, is it? Glomb?

Member Glomb: Hi, I'm Brendle Glomb, Senior Medical Director for Superior Health Plan. I oversee value-based contracting and our quality programming. I have nothing to disclose.

Dr. Pickering: Great. Thank you so much. And before going to our surgery standing committee colleagues, I'm just going to circle back to James. James, are you able to get off mute?

Member Rosenzweig: Yes, I'm Jamie Rosenzweig. I'm an endocrinologist with Hebrew Senior Life and with the Endocrine Society.

And I have nothing to disclose except that I was involved in developing measures that we were discussing 20 years ago.

Dr. Pickering: Great. Thank you so much, James, for the disclosure. I appreciate you being able to get off mute.

Okay, so just checking in again. Do we have Anna McCollister or Carlos Bagley on the line? Okay.

So I'm going to go to our Surgery Standing Committee members. So welcome and thank you very much for your participation in today's proceedings.

What I have listed here on the slide, I'll just go

down the list. The same applies to you all as well as just the name, your affiliation as far as your organization, and also any disclosures you'd like to list. So Vilma Joseph?

Dr. Joseph: Hi, yes, I'm Vilma Joseph. I work at Montefiore Medical Center, Albert Einstein College of Medicine, and I am associated with the American Society of Anesthesiologists. And I have nothing to disclose.

Dr. Pickering: Thank you so much, Vilma. And Richard, and I apologize about this, D'Agostino?

Dr. D'Agostino: It's Richard D'Agostino. Hi, I'm a cardiac surgeon at Lahey Hospital and Medical Center. I have nothing to disclose related to the measures we're discussing today.

I am a member on the Adult Cardiac Surgery Database Task Force for STS. And I do participate in a working group that looks at the results generated by the pre-operative beta blocker measure.

Dr. Pickering: Okay. Thank you very much for that, Richard, for those disclosures. Miklos Kertai? Is that correct?

Dr. Kertai: Thank you. Good morning, everybody. My name is Miklos Kertai. I am a cardiothoracic anesthesiologist at Vanderbilt University Medical Center.

I have nothing to disclose with regards to the measures that will be discussed today. But I am a member of the Society of Thoracic Surgeons throughout the Quality Improvement Committee. Thank you.

Dr. Pickering: Great. Thank you so much. Michael Firstenberg?

Dr. Firstenberg: Good morning, everybody. My name is Michael Firstenberg. I'm an adult cardiothoracic surgeon.

I'm actually in between positions right now, but I'd prefer for confidentiality reasons not to disclose my current employer that I'm working with.

I'm a member of the SPS. I've been involved in various committees and consulting work over the past five years, but nothing pertaining to any of the topics that we're talking about today.

Dr. Pickering: Great. Thank you so much, Michael. And lastly, Salvatore Scali? Salvatore Scali? Okay. And just one last time, Anna McCollister or Carlos Bagley?

Okay. So thank you all very much. I'd like to let you know that if you believe that you might have a conflict of interest at any time during the meeting as topics are discussed, please speak up.

You may do so in real time during the led meeting or you can send a message through the chat to our chairs or to anyone on NQF staff. If you believe that a fellow committee member may have a conflict of interest or is behaving in a biased manner, you may point this out during the meeting, send a message to the chairs, or to NQF staff as well.

Does anyone have any questions or anything you'd like to discuss based on any of the disclosures that have been made today?

Okay. Nothing in the chat. No hands raised. Okay. Thank you very much. And lastly, as a reminder, NQF is a non-partisan organization.

Out of mutual respect for each other, we kindly encourage that you make an effort to refrain from making comments, innuendos, or humor relating to, for example, race, gender, politics, or topics that otherwise may be considered inappropriate during the meeting.

While we encourage discussions that are open, constructive, and collaborative, let's all be mindful

of how our language and opinions may be perceived by others.

With that, I will turn it back to the team. So thank you all once again very much, and looking forward to the rest of the proceedings.

Overview of Evaluation Process and Voting Process

Ms. White: Wonderful. Thank you so much, Matt. We do appreciate you joining us on the call today.

So I'm going to actually turn it over to Isaac Sakyi, our team manager, who will present and overview of the evaluation and voting process.

So if we can go to the next slide, please. All right, so, Isaac.

Mr. Sakyi: Can we go to the next slide? So, the next one. Thank you. I'll be going over the evaluation process that will be followed today.

Our Standing Committee members acted the proxy for the NQF stakeholder membership. They evaluate each measure against each criterion and with that indicate the extent to which each criterion is met and the rationale for the rating.

They also respond to comments submitted during the public commenting period, make recommendations regarding endorsements to NQF members, and oversee the portfolio of PCCI measures. Next slide.

To go over some ground rules, we would like to emphasize that this is a shared space and there's no rank in the room.

We encourage you to remain engaged in the discussion without distraction, and hope you are prepared and have already reviewed the measures.

Please base your evaluation and recommendations on the measure evaluation criteria and guidance.

Keep your comments concise and focused, be cognizant of others, and make space for others to contribute to the conversation. Next slide.

In terms of how discussion will proceed, we'll start with a brief introduction of the measure by the measure developer.

The lead discussant will then briefly explain the information provided by the developer on each criterion, followed by a brief summary of the evaluation comments from the committee, emphasizing areas of concern or differences of opinion.

The lead discussants will also note the preliminary ratings by NQF staff, which is intended to be used as a guide to facilitate discussion.

Developers will be available to respond to questions from the Standing Committee. Afterwards, the full Standing Committee will discuss, vote on the criterion if needed, and move onto the next criterion. Next slide.

The following is a list of our endorsement criteria. Five areas are outlined here, mainly importance to measure and report, which includes evidence and performance gap; scientific acceptability, which also includes liability and validity. Please note that the first two bullet points are a must pass criteria.

We also have feasibility, usability, and use, and related all competing measures. The use sub-criterion is a must pass for maintenance measures.

The next point of discussion is the comparison to related or competing measures, which is a discussion and does not require a vote.

A discussion only takes place if the measure is recommended for endorsement. Next slide.

So again, these are the criteria the measures are evaluated and voted on. Next slide.

If the measure fails on one of the must pass criteria, there's no further discussion or voting on the subsequent criteria for that measure.

The committee's discussion will move on to the next measure, if applicable. In our case, we have multiple measures to ensure steps in the process should we find ourselves in the case where a measure does not pass on a must pass criteria.

If consensus is not reached on a criterion, the discussion will continue on to the next criterion.

But ultimately, there will not be a vote on the overall suitability for endorsement. Next slide, please.

As far as achieving consensus, quorum is 66 percent of active standing committee members, and that is 13 members for our 19 active standing committee members.

We need greater than 60 percent of those to pass the criterion or recommend a measure for endorsement.

So, yes votes are a total of high and moderate votes. Between 40 to 60 percent of community members voting yes will be consensus not reached, and less than 40 percent voting yes means the criterion does not pass or the measure is not recommended, depending on what we're voting on.

The consensus not reached criterion and a vote of overall suitability for endorsement would be postponed to the post comment meeting.

If a measure is not recommended, it will also move on to the public and NQF member comment, but the committee will not revote on the measure during the post comment meeting unless the standing committee decides to reconsider based on submitted comments or if the developer submits a formal reconsideration request. Next slide.

As mentioned before, please let us know if you need to step out of the meeting. We need quorums to vote on the measures, and that means 50 percent of the standing committee on the call to continue the discussion.

If we lose quorum at any point, we will shift to an offline survey, which will contain the same questions as the live voting platform.

In that situation, we will ask that the standing committee submit their votes within 48 hours of receiving the survey and the transcripts of the meeting.

If a standing committee member has to leave and we still have quorum, the committee will continue with the vote and the discussion.

The standing committee member who left will not have the opportunity to vote on the measure evaluated during their absence.

That sums up the process for today's meeting. And at this point, I would like to pause to see if there are any questions.

Co-Chair Bratzler: Yes, Isaac, have we done a count? So do we know what our quorum is for today? I know we had a couple of members that couldn't be here.

Mr. Sakyi: Yes, so quorum today is 13 members, and we have 16 standing committee members on the call.

Co-Chair Bratzler: Okay. Very good.

Mr. Sakyi: So hearing no other questions, I will turn it over to Tristan for a voting test.

Mr. Wind: Thank you, Isaac.

Ms. White: And while Tristan is pulling up his slides, again, please let us know if you have not received

the link in your email. We will be happy to work with you to get that. Okay, Tristan.

Voting Test

Mr. Wind: Thank you, LeeAnn. So good morning and thank you for attending today's call. We sent a voting link via email.

If you did not receive that, please let a team member know and we will be happy to assist.

So today's test question is have you visited the beach? Select A for yes and B for no.

Dr. Firstenberg: This is Michael Firstenberg. I'm listening in on my phone but I have to step away from my computer for a few minutes.

So I'm not going to be able to vote for a little bit, for about 15 to 20 minutes if that's okay.

Ms. White: Michael, that's perfectly fine. And if you would also, if you would directly message the team with your vote, we can also work with you to do that as well.

Mr. Wind: Thank you, Michael.

Dr. Firstenberg: Okay. Thank you.

Mr. Wind: So for today's purpose for voting for the test, we will have 15 votes instead of 16.

Member Glomb: This is Brendle Glomb. I'm just not getting a response on the screen. Where?

Mr. Wind: So using the voting link, so you will be prompted to type your name and then that will pull up the question.

Member Glomb: Sorry about that.

Mr. Wind: No, you're okay.

Member Glomb: Did we lose everybody? You guys

still there?

Mr. Wind: We're here.

Ms. White: Yes, we're here. We're just waiting for one more vote. Okay, we're at 15.

Mr. Wind: Perfect. So that conducts today's voting test. 53 percent voted yes, 47 percent voted no.

Ms. White: In all fairness, it's the beginning of summer, so we all have time to get to the beach. Okay. Perfect. All right, thank you, Tristan.

All right, so, we are going to pull back up our slide deck here to start our measures under review. One moment.

The logistics of virtual meetings here. Victoria, I think you need to share your screen again. Perfect. Okay.

So I will go through measures under review for this cycle. Next slide, please. And just provide a brief overview.

So we received three maintenance measures and one new measure for the Spring 2022 PCCI cycle.

The three maintenance measures are 0729, optimal diabetes care, 2797, transcranial doppler ultrasonography screening among children with sickle cell anemia, 3294, SPS lobectomy for lung cancer composite score, and the one new measure for standing committee review is 3668, follow up after emergency department visits for asthma. Next slide, please.

We did not have any measures go through SMP review this cycle, but we do like to talk about the SMP and the role in the measure review process.

The Scientific Method Panel is a group of researchers, experts, and methodologists in the healthcare quality measurement arena.

The panel reviews complex measures and provides comments and concerns to the developer. The developer than has the opportunity to provide further clarification and update their measure submission form before the standing committee evaluation.

Again, no measures were reviewed by the SMP for the Spring 2022 cycle. Next slide, please. Great. And then next slide, please.

So now we have reached the part of our measure evaluation meeting where we begin the consideration of our candidate measures. We'll go to the next slide.

So with that, our co-chairs will begin by introducing the measure. The developer will then have the opportunity to provide a 3--5-minute overview of their measure.

Our lead discussants will introduce the criterion and highlight their main takeaways. Our supporting discussants will respond to the lead discussant and add their insights.

During the criterion discussion, the co-chairs and staff will collect those questions for the developer.

Once the initial discussion on the criterion is complete, the co-chair will then ask the developers to respond to questions and then clarify any information.

Once the standing committee has completed its discussions, a vote will be taken on the discussed criterion.

3668: Follow-Up After Emergency Department Visits for Asthma

So our first measure is Measure 3668, Follow Up After Emergency Department Visits for Asthma.

The measure is at Albert Einstein College of

Medicine. And the developer is UCSF. This is a new measure.

The brief measure description, it is a process measure that seeks to capture follow up after asthma-related emergency department visits for children with asthma after discharge from the emergency department as recommended by NLBI 2007 guidelines.

The measure also assesses the percent of asthma-related ED visits for children ages 3-21 with a follow up visit with a primary care clinician or an asthma subspecialist within 14 days of discharge from the emergency department.

Within the reporting year, the patients who are enrolled in the health plan for two consecutive months following the ED visit.

So with that, I will pass the baton over to Dr. Dale Bratzler, who will take on this discussion. So, Dale?

Co-Chair Bratzler: Yes, thank you, LeeAnn. So this is a new measure, so we'll be going through a complete evaluation of this particular measure.

And I believe we have a representative of the developer here. Is Naomi on the call? Do we have somebody from the developer on the call? I can't see everybody on there.

I know it's early in San Francisco. LeeAnn, do you know if the developer is on the call?

Ms. White: I do not. I have some call-in users. I'm not sure who those are. And I will also look to our inbox to make sure to check those communications.

Co-Chair Bratzler: Well, hearing nobody from the developer at the moment, should we just go ahead and start the conversation?

Then if the developer comes on the line or notifies you, we will stop and give them a few minutes to

describe the measure.

So Anna had notified us that she wouldn't be able to be with us first thing this morning.

So I believe Starlin has agreed to be the lead discussant for this particular measure. So Starlin, I'm going to turn the conversation over to you at this point.

Member Haydon-Greatting: Okay. Yes. So as what was previously read, it is a new measure, follow up to emergency department visits for asthma in ages 3-21.

It's a process measure. It uses Medicaid claims data. And it's important to note that Medicaid claims data is sometimes very different than commercial insurance claims data.

We're trying to make it all the same, but we probably need to make note that it is a Medicaid claims data.

And the level of analysis is for a health plan. And so with their evidence, they used a systematic review.

One of the authors of the systematic review is part of the review part of the team that developed and did the measure.

And I'll go right to some of our committee comments. My first question is because of the hats that I wear, is sometimes California defines eligibility for Medicaid up to 21 years of age and then they've added up to 26 years of age in this past COVID.

So that's important to know. In Illinois, we cut it off at 18. So each state sets their age standards for pediatric care that's being covered under the Medicaid.

If they still require Medicaid after the age of 18, they move into the adult phase. So that was a

clarification I was looking for.

So this process measure, the evidence from 2007 National Asthma Education and Prevention Program Guidelines for the Diagnosis and Management of Asthma Summary Report.

They have a new 2020 report, but the 2020 report dealt mostly with medication updates. I reviewed both of those reports.

And the guidelines were -- and they state that they were updated but it didn't impact the works of follow up.

Now I completely support a follow up because, as we know, specifically in the Medicaid population, sometimes they don't have a primary care professional.

And the emergency room is used as their primary care delivery system. So the evidence is there for that.

For the gap in care, they did an excellent review of their performance, but in my comments, I did point out that the 18-to-21-year-old area didn't perform as well as the 3-to-18-year-old group.

And in my personal experience, we have fall out. We stop the patients after 18 and they may not stay in the Medicaid or any insurance system.

Or they may switch to a different insurance. Some colleges provide insurance, so forth and so on. But in their data set, they did an excellent job of that point.

The lowest dectile was 11.7 percent. And with an NNI health plan, this comes from their systematic review, which also had high marks for the systematic review that they used for their evidence.

Performance in the highest dectile was at 43 percent. And overall performance was 22.1 percent.

The follow up visits were higher in patients ages 3-5 and those who were Asian Pacific Islanders had a higher percentage of follow up visits compared to those who were Black, Hispanic, and White.

Some of that may be the geography. So California has many more people of those diversified cultures.

So, when you apply those to other populations, you may need to expect that culture shift may happen.

Patients with Medicaid fee for service had a 10.2 percent, and were less likely to have a follow up visit compared to those with managed Medicaid, which was 23.9.

Now, I need to also let you know that managed Medicaid in their contracting to cross off that they have that follow up visit.

So this is a good performance measure to ensure that that is happening. Do you want me -- we need to stop after evidence, right?

Co-Chair Bratzler: Yes, I think so. So, I had one question for you, Starling. What --

Member Haydon-Greatting: Okay.

Co-Chair Bratzler: I know the guidelines recommend the follow up and the performance rates listed here were quite low.

Member Haydon-Greatting: Yes.

Co-Chair Bratzler: Is there good data on changes in patient outcome?

Member Haydon-Greatting: So, I think there is, but this is a Medicaid population and they mostly referred to their systematic review.

I tried to pull some of the articles out of that systematic review that they used to support that. And it was good support. Oddly, or serendipitously,

JAMA just published an article yesterday that said a phone call within 2-14 days after emergency room visits improved all patient outcomes, no matter if it's asthma or not.

And in my own personal experience in doing chronic care conditions, having follow ups, getting connected to a health professional post their emergency visit, it leads to good outcomes in the evidence. And their systematic review did include many of those articles that support that.

Co-Chair Bratzler: Okay, thank you. So before we go on at this point, I'm going to see, Lindsay, you were the secondary reviewer. Do you have any particular comments on the evidence for this metric?

Member Botsford: Yes, thanks, Dale. Yes, I think similar to Starlin, they do have a combination of a few small randomized central trials and then some observational studies that do confirm the association between follow up visits and decreased ER utilization in the future.

What is a little bit unclear is you don't know what the content of the visit was at follow up and was there anything different about the content of the visit?

Was there an asthma action plan done? Or was this visit actually happened to not discuss the asthma but happened in a PTP setting in that timeframe?

So I think not complete homogeneity in terms of what happens at these visits that would be a visit.

So it does seem like there's a strong association. What was less clear to me if this was causation or other factors that would encourage or that would enable a patient that would have a follow up visit, also decrease the likelihood of future ER utilization.

The other thing I think that was a question for the

developer, if at some point they join us, is it looks like the 14 days was chosen.

In the 2007 NHLBI guidance, it says the recommendation is that the follow up happen one to four weeks with a PTP or asthma specialist.

Now, since those guidelines came out, there have been studies that anchored on 14 days, which supports their choice of 14 days, but I would question that.

And then the final question I had was getting towards what the most recent discussion was on, is what type of visit counts?

In a world where virtual and phone encounters exist, are those included in the claims? Or is this an in-office visit only. And that was potentially shortsighted.

Overall, I think while the number needed to treat seemed a little bit high to me, the potential reduction of half the population does feel real safe.

We'll get into some of the disparities in discussion later, but those of you that constantly think about evidence it would be helpful to probe a little bit more.

Member Haydon-Greatting: Yes, and the study pulled claims from 2015 and so we were -- while California is ahead of the game in the telehealth telemedicine world, we don't know if that was utilized.

I know that California Medicaid was part of some demonstration projects in their rural areas to enhance telemedicine for chronic conditions, asthma, COPD, diabetes, and cardiovascular disease.

But again, they did not mention any of that. If that is just the fact that, I know that because I'm in that world. But that would be helpful to know. So I

concur.

Co-Chair Bratzler: All right. Thank you both. So I'm actually taking notes in case we get the developer on the call so we can summarize those questions.

Anyone else on the committee have any comments about evidence? We'll limit our discussion to evidence at the moment.

Member Rosenzweig: Excuse me. How are the follow up visits documented? If a person, let's say, moves out of the state or something to a different area, and does a telephone call, is that considered a follow up?

Co-Chair Bratzler: Yes, I think that was one of the questions that Starlin and Lindsay both have is telehealth, phone call follow up, do any of those count?

Since Medicaid is claiming, I'm assuming there would have to be some type of encounter documented in claims.

Member Haydon-Greatting: So that goes along with my question. So in the claims data, we have G codes that indicate whether it was audio or telemedicine that go along with that Medicaid claim that could be pulled at the same time when they're doing their preliminary pull.

The other concern I had was during routine patients that had a primary care provider prior to being looked at.

So if you don't have a primary care, established primary care provider, you will use these emergency rooms.

So when I did a similar study like this in Illinois in the Chicago area, and we went through and pulled everybody who had a primary care provider attached to them, then did the same analysis to see what was happening in the emergency room,

recidivation visits.

You first need to establish a primary care provider's in place. Then you pull that data. Then you can look and see what kind of post data visits happened after the emergency DRG code. The other concern that I have is depending on the state Medicaid, sometimes the state Medicaid grants hide Medicaid to hospitals, blanket money to take care, because they use a primary care provider in an underserved area. And so you're not going to see those in that data. So in the evidence, it would be nice to kind of get a better idea of, I mean, I'm not suggesting they give us all the claim codes because that would make us dizzy.

But if there was a G code you could tell if it was a telehealth follow up visit, that sort of thing. Anyway.

Co-Chair Bratzler: All right. So, I mean, I think we're circling around. We have a question or two here for the developer about how this was accounted.

Starlin raised some issues. I know I read your comments earlier about does the patient have a pre-existing PCP because it is urgent to maybe get into one if you don't have one with you. So any other comments about evidence? And we'll move to that vote.

Member Glomb: Dale, this is Brendle Glomb.

Co-Chair Bratzler: Yes.

Member Glomb: I'm a pediatric pulmonologist. I guess one of the things that I'm concerned with is again returning to the data itself.

In our data, Texas statewide, which is PCP Medicaid, we've got a lot of provider turnover diagnosis of asthma in the emergency room, first time wheezers, infants, young toddlers, as an example, with bronchiolitis.

Is there a plan within this data analysis to sort that

out? As an example, this patient may be well known to the PCP, but for whatever reason, however, the patient winds up in the emergency room, has the tag of asthma, but on that visit then they're back to the PCP who knows this is -- knowing this patient, knows that this is not asthma in the patient.

There's no way to filter out some of this. If it's not really asthma, if a PCP knows the patient well, is there a reason for it?

And I'm being, yes, I think there is, personally, but is there a reason for having a visit within the 14-day window? I hope I'm clear on that.

Co-Chair Bratzler: Yes, thank you. Ann, you had your hand up.

Member Kearns: Yes, I'm not sure if my question is relevant to evidence or not but I'll raise it, and that is if the ER visit happens in a system different from the primary care, how does the primary care -- and maybe I don't know enough about how Medicaid functions to alert a primary care, but in the adult world where I see people and they don't know who their primary care is or it's a team and you're trying to alert them to they're not in my system, and then you have such a short window of 14 days to communicate with the primary care to get the appointment set up, it just seems, maybe that's more a feasibility issue and maybe this is not an evidence issue. So tell me to hold it if that's the case.

Co-Chair Bratzler: All right. Thank you. Also I'm going to go to one more person. Curry, you had your hand up. Then I believe the developer's joined us so we'll break and let him give a brief introduction.

Member Curry: Thanks, Dale. There is some evidence in the Medicaid population that a visit with a primary care provider may not impact the ED readmission rate or utilization.

However, if that's done in primary care based, as in special equipments, there's additional, significant reductions in utilization.

So I guess the question is, from the outcomes, does it matter where they get their follow up? And there might be some evidence that says no.

Co-Chair Bratzler: All right. Thank you. Naomi, I understand that you've joined the call from UCSF?

Ms. Bardach:: Hi, yes, good morning. Apologies. I'm coming from another NIH review committee and I thought we were starting at 9:55 with the measures. So I apologize for being a little bit late.

Co-Chair Bratzler: Okay. So if you would, we'll go back, start to give this 3-to-5-minute overview of the measure, and then a number of questions have been answered.

Ms. Bardach:: Excellent. So this is the measure. It is a claims-based measure focused particularly on pediatric asthma care, which many of you know is a highly prevalent condition in pediatric patients than leads to paramedic healthcare utilization as well as missed school days and work days for families.

The measure itself is looking at, for those patients who go to the emergency department for asthma-related visit, for children ages 3-21, the question is whether or not they have a follow up visit.

So the measure assesses whether or not they had a follow up visit. So the measure assesses whether or not they had a follow-up visit with another primary care physician or an asthma subspecialist within 14 days of discharge from the emergency department, for patients who are enrolled in the health plan for at least two consecutive ED visits.

And it's an administrative claims-based measure. And we tested it in the Medicaid population in California.

We also published a paper looking at it in Massachusetts and Vermont data as well, which we didn't include in this particular submission, but that's the sort of evidence behind it in addition to the fact that it's consistent with NHLBI guidelines for asthma care for pediatric patients.

Co-Chair Bratzler: All right. Thank you, Naomi. So I'm going to try to summarize a few of the questions that came up at our discussion of evidence.

One thing, because this is an administrative claims-based metric, what about follow up via telehealth or phone call or other follow up? How is that handled in the metric?

Ms. Bardach:: That's a great question. It's anything that is a claim in the administrative claims-based data, so it's all, it's going to be Medicaid data or commercial claims data.

So anything that's claimed is going to be included in the measure as well as it has the correct IGB code and the provider type is captured by the measure specifications.

So it's a great question because we did our data analysis, you know, we used data that was before the pandemic.

And so telehealth was just not going to have been as common a mechanism. So it really gets to the question of, in the regular world of who's going to be reimbursing for telephone, whether or not those telehealth visits are actually going to be counted in a claim or not.

Right now, the measure specifications do not differentiate between a claim and, oh, health versus a claim that is not telehealth.

Co-Chair Bratzler: All right. Thank you. Then another question came up about just diagnostic

accuracy.

So perhaps the patient has been seeing a primary care physician over time, gets a diagnosis of asthma in the emergency department, where there are some concerns about diagnostic accuracy.

The family physician, of course, primary care or whoever is taking care of them, knows that he actually doesn't have asthma, but that diagnosis comes up.

Ms. Bardach:: And so the question is whether it would be appropriate for a follow up visit, meaning like the PCP might then say, oh, you don't actually have to come back in?

Co-Chair Bratzler: Yes.

Ms. Bardach:: I would probably argue that clinically that might not be actually a practice, because if you have a patient who is in the emergency department and is diagnosed with asthma for the first time, I would suspect most clinicians would not say, oh, that person should not come in because I know for sure they don't have asthma.

I think that would be actually a situation where kind of you hear a clinician say, oh, that's new, we should make sure that the kid is doing okay."

But I understand the concern. I think that that might not actually be something that happens in practice.

And it's also a health plan level measure, so it's a question of, it's not the common cure physicians who would be then the account entity who would then be considered underperforming because of the health plan measure.

Co-Chair Bratzler: Okay. And then the last part, and I'm sure I thought was very interesting.

Have you stratified the measure based on whether

the patient had a documented assigned PCP before the event?

In other words, somebody who presents for the first time to the ER with diagnosis of asthma who doesn't have preestablished PCP.

It may be quite difficult to get in to a PCP with that two weeks. Have you stratified the data that way?

Ms. Bardach:: So we looked at what we termed identifiable asthma versus not identifiable asthma, and the identifiable asthma is people who have had a diagnosis of asthma prior, looking at their one-year looked at period in their claims data.

So that. And we found that the follow up visit was associated with decreased utilization, subsequent decreased utilization, whether you have identifiable asthma or not identifiable asthma.

It tends to actually be a little bit more, a strong confusion if you didn't have identifiable asthma, I think in part because of the fact that then those patients got connected.

If they were connected, then that was actually a good sign. And they were able to then have their asthma managed.

And so we did not do a stratification by whether you had a PCP prior or not. But it does look like it's probably associated.

But there are some association between people who have had a PCP prior or not, because we also see a pretty long-term effect all the way out to a year of having that following.

But it's have an indicator not only of that one visit, but actually, that you have an established primary care physician, meaning that you already were able to get into that PCP relatively easily.

Therefore, you have that visit. And therefore, you

probably have a relationship that then leads to ongoing monitoring and treatment of your asthma a little bit more effectively.

So, I do think there are signs signaling, but actually I would say because of the health plan measure, that sort of appropriates that health plan should be helping to facilitate people in getting connected in PCPs and stop having to the primary care environments that they have access to primary care physicians.

Co-Chair Bratzler: Yes, thank you for continually reminding us, it's at the level of the health plan. That's helpful. Adam --

Member Glomb: This is Brendle. Can I ask, can I insert a question with regard to that health plan evaluation? Sorry.

I'm sorry, I'm a pediatric pulmonologist with a same team corporation. Given the often-lengthy claims submission, how is that dealt with or is it dealt with within the measure if the follow up is supposed to occur within this window, but claims submission, because this is a claim based, claims based measure, is not in an almost instantaneous fashion, how would you suggested then that the plans deal with that lack and their ability to point the patient into their PCP?

Second to that, in Texas, all children in Medicaid are required to have a designated PCP. Whether or not they utilize that PCP or not is a different story. But they all do have one.

So, I'm just wondering about the claims lag and then that responding to the plan, who may not even know for 14 days. This is an example that such a visit is good.

Ms. Bardach:: Yes, great question. I read that the way I think health care organization and how a health plan might have a level of control over this

would actually not be that it's on them to call the patient as soon as they go to the emergency room, but actually that it would be something that they would create with care, exactly to your point where there's a requirement that they be connected to have a PCP and that there's accidents PCP appointments.

So, and that goes to other part measures, like available appointment in the primary care setting, but I don't think, I think your point is well taken, but I don't think the measure would exist in a vacuum, and in general, we should be moving towards getting better connection to the PCPs or to care after the -- at the emergency room visit.

I will make one more night, which is Dr. Cabana is on the measure -- he's the measure steward.

And the measure co-developer is here with us also. We were in the same NIH review section. He is coming over from that meeting as well.

Co-Chair Bratzler: Thank you. Adam, have I missed anything?

Co-Chair Thompson: I think the only lingering question, and it was somewhat I think even resolved in our committee discussion but I want to make sure we put it out there was the difference between the recommendation being followed up within one to four weeks and the measure using 14 days and whether there was any specific reasoning behind that decision.

Ms. Bardach:: You'd have to, the 14 day window, we started at different windows, a 7-day window, 14-day window, 30-day window. The association was that the win was the strongest when it was within the 14-day follow up and not visits within 7 days, to actually be able to get a whole lot of signa because you just, methodology is so low.

14 days was a stronger association of having

decreased reutilization subsequently compared to a 30-day window.

Member Glomb: Great.

Co-Chair Bratzler: Thank you. So I think at this point, unless there's anybody that has things in the chat, I think we've captured all those and we'll go ahead and take a vote on evidence.

Ms. White: Okay. So if you'll just bear with us just a moment, Tristan will bring up his Vote Everywhere screen.

And if you just recently joined us, I know we had Carlos join us, Carlos, welcome onto the call. We do just need to quickly go over a quick introduction.

So, if you can please just put your name and your organization and any disclosures that you may have before we proceed to the vote, I'd greatly appreciate it.

Carlos, if you're able to come off mute. I will go ahead and unmute you. We will work with Carlos behind the scenes. So, okay, Tristan, go ahead, and if you could pull that up, please?

Mr. Wind: Thank you, LeeAnn. So voting is now open for Measure 3668 on evidence. The options are A for high, B for moderate, C for low, and D for insufficient. As a reminder, we will be waiting for 17 votes.

Member Rosenzweig: Excuse me, I am clicking on the boxes and nothing shows up. Am I on the wrong screen?

Co-Chair Bratzler: Yes, don't use the slides that they're showing on the screen. You have to go to the link.

Member Rosenzweig: Sorry.

Co-Chair Bratzler: I tried the slide, too, the first

time. Then I realized I was looking at the slide.

Mr. Wind: That's a great point. As a reminder, we will be using the same exact link for the duration of this meeting.

So again, we will be waiting for 17 votes. Is anyone having any difficulty right now? Because we're awaiting two more votes. So please raise your hand or contact a team member in the chat.

Co-Chair Bratzler: James, I see your hand up.

Ms. White: If you are unable to vote using the link, please just message Isaac Sakyi, our manager, directly.

So, if you go to the chat box, you can drop down, find his name, and you can message us privately your vote.

Mr. Wind: And last call for voting. Okay, it looks like we have received all the votes. Thank you for your patience.

So voting, we received 0 votes for high, 14 votes for moderate, 2 votes for low, and 0 votes for insufficient.

Please provide the team one moment to confirm the votes. Therefore, the measure passes on evidence. Thank you. Back to you.

Co-Chair Bratzler: All right. Thanks. So if we will, we'll go back to Starlin, and if you'll talk about performance gaps.

Member Haydon-Greatting: Okay. So in the performance gap, the lowest dectile is 11.7 percent and performance in the highest dectile was 43 percent.

Overall performance is 22.1 percent. And the high follow up visits were higher in patients in ages 3-5 as we mentioned before, and lowest in patients 18-

21.

And then I pointed out previously that the cultural differences, Asian and Pacific Islanders had a high percentage of follow-up visit as compared to those who are Black, Hispanic, and White.

And the fee for service Medicaid were less likely to follow up as compared to the Medicaid and the managed care group.

Now, the managed care group, this is a requirement in most managed Medicaid contracting. So they have a moderate performance gap. And Lindsay, do you want to add anything to that?

Member Botsford: No, I mean, I think the evidence highlights the gap. Some of it is bigger in certain populations, so we can get into some of the conversations later.

Co-Chair Bratzler: Yes.

Member Botsford: But the performance, that seems fairly good evidence that it exists.

Co-Chair Bratzler: All right. Thank you. So anyone else on the committee have any comments about performance gaps?

So, Adam, I'm counting on you to keep me honest about hand raising because I don't see those all the time.

So if there are no other comments at. His point, I think we could go ahead and look at performance gaps.

Mr. Wind: So voting is now open for Measure 3668 on performance gap. The options are A for high, B for moderate, C for low, D for insufficient. And we are waiting for 17 votes.

Ms. White: And while we're voting, I just want to try again to reach out to Carlos Bagley.

If you can please come off mute and introduce yourself, your organization, and any disclosures that you may have.

Mr. Wind: Okay, voting is now closed on performance gap for Measure 3668. We received four votes for high, 12 votes for moderate, zero for votes for low, and zero votes for insufficient.

Mr. Sakyi: Just a correction, we have one more vote, so 13 votes for moderate.

Mr. Wind: Thank you, Isaac. Please provide the team one moment to confirm the votes. Measure passes on performance gap. Thank you. Back to you, LeeAnn.

Co-Chair Bratzler: All right, thank you. Starlin, if you would, go ahead and start talking about scientific acceptability. We'll focus on reliability first.

Member Haydon-Greatting: So the reliability specification, I need to point out that each state's Medicaid license administers their own program and it can be different state by state.

And so whether for fee for service or managed Medicaid, the only thing that is somewhat consistent across all 50 states and territories is the managed Medicaid guidelines, or for them to go into place before the CMS accepts the managed Medicaid profile for a state.

So if this goes to use for a health plan, the health plan will have to -- so if they're -- if they're doing a managed Medicaid program, they're going to have to assess what their state requirements are.

If it goes to commercial, some of the same thing. So there are state to state differences, as we all know, in this world.

The Medicaid data limits the ability to do crossover into a commercial plan to some extent.

I still think it's a -- I still believe this is a useful measure because we don't do enough follow up with children, but I'm just saying that in my 23 years of Medicaid practice, I was -- I was never able to go because I was also -- I also did the employer/employee health for Illinois and I couldn't use -- I couldn't use one set of data to make the commercial ensure people use the same measure of quality. So I just want to bring that up.

But their specifications are clear and precise. And they did -- they did -- they did very well consistently implement that reliability.

They had great quartiles and great confidence intervals for their random split half reliability testing.

And they were able to compare the fee for service with the managed Medicaid, which was -- which was very well done.

Co-Chair Bratzler: All right, Lindsay, do you have any comments?

Mr. Sakyi: I think the only other thing to build is calling out it did look at over 100 plans and had some exclusions for plans with a small number of eligible patients, which seemed to account for issues with plans that might be skewed by just having a smaller population.

But they did account for some thoughtful things and in the testing. So again, maybe some concerns to talk about later, but in reliability, I think no major concerns noted here.

Co-Chair Bratzler: All right, thank you, Lindsay. Any other committee members that have comments about reliability?

All right, hearing none, I guess we can move to the vote on reliability.

Mr. Wind: Voting is now open for Measure 3668 on

reliability. The options are A for high, B for moderate, C for low, D for insufficient. And we will be waiting for 17 votes.

Voting is now closed for Measure 3668 on reliability. There were six votes for high, nine votes for moderate, one vote for low, and zero votes for insufficient.

Please give the team a moment to confirm the votes. Therefore, the measure passes on reliability. Back to the co-chairs. Thank you.

Co-Chair Bratzler: Thank you. Starlin, I'll turn it back over to you.

Ms. White: Sorry, I just need to make a correct to the record, Dale, real quick. So we need to correct the record for the gap.

And we did have someone vote twice. So we just need to make sure for the record, we have for the results on performance gap, we had four votes for high and 12 votes for moderate.

The gap criteria still passes with 16 votes for moderate and high, but we went down to 12 from 13 for moderate. So I just need to make that quick correction. Thank you.

Co-Chair Bratzler: All right. Thanks for keeping us honest here. All right, Starlin, talk about validity.

Member Haydon-Greatting: All right. So they did two validity testing. They did a patient encounter level and they did it at the -- also tested at the accountable institute level.

And their model was adjusted for average age during the index year, gender, chronic disease status, and insurance type, evidence, and making sure that the evidence was that they had asthma.

They used a strata post estimation margin, which shows up with patients at the follow up within their

14-day window and in that 365 days of asthma related ED visits.

And they showed the lower rates of subsequent asthma-related utilization as 5.7 percent with a good confidence interval and also 5.7 percent to 25 percent as compared to no follow up, which was 6.4 percent to 28.3 percent.

The validity on the accountable entity level, they assessed the relationship between performance on the measure and the eligible patients and the repeat utilization.

They had a beta quote efficient of negative 0.19 with a p value less than .001 for 50-day revisits, and for each 1 percent increase in follow up visits, there was a decrease of 0.2 percent and a 60-day emergency room revisit.

And the staff rated this as moderate for ability.

Co-Chair Bratzler: All right. Thank you, Starlin. Lindsay?

Member Botsford: Yes, I did find this part interesting in that it sounds like some of the previous literature 20 years ago, looking at the correlation between ER utilization and PCP visits, actually didn't show an improvement, or if anything, showed a paradoxical increase in utilization, perhaps due to bias and selection.

I think it sounds like the results of the testing are currently in craft. So I don't know if there are any updates on that.

So also, it sounds like the claim sets that were used were more than one state. I think also Massachusetts and Vermont, so we're not just looking at California data and looking at whether there are correlations.

So did seem to be statistical significance in those populations showing performance on the quality

measure.

So comments or concerns I think. I did still look like again a thoughtful approach to confirming the association in a variety of data sets, especially in light of previous evidence that would suggest not a benefit to this process measure.

Co-Chair Bratzler: All right. Thank you, Lindsay. Any other committee members that have comments about validity?

Dr. Joseph: Hi, yes, this is Vilma Joseph. I just had a quick question regarding the fact that there was no missing data at all.

I'm just amazed at that. So if the developer could explain that, or was it a significant amount of data?

Ms. Bardach:: Yes, thank you so much. I was reading the review comments yesterday. And so it was just very small numbers of the data elements.

So I had our data analyst pull some of our data for that just to get a general sense for the key elements.

So for age, we had, it was such small percentages that we just didn't delve deeper into doing analysis of it.

It was 0.00004 percent, so very few missing values for age in the California Medicaid data set, and then larger missing data for diagnosis number one.

It was still very low, just 1 percent, max of 2.3 percent for 2015 data. We looked at 2014, 2015, and 2016 data, and it was 1.2 percent for the 2014, 1.4 percent for 2015, and then 2.3 for 2015.

And then for the other piece that would be, the measure was the NPI number, meaning how do you identify which provider was associated with the claim.

And we again had very small numbers, 0.003 percent to even lower than that across the years. So very little missing data.

Dr. Joseph: Okay, thanks.

Co-Chair Bratzler: Other committee members with questions or comments about validity? All right, I think we can move to the vote on validity.

Mr. Wind: Voting is now open for Measure 3668 on validity. Options are A for high, B for moderate, C for low, and D for insufficient. We will be waiting for 16 votes.

Voting is now closed for Measure 3668 on validity. There was one vote for high, 14 votes for moderate, one vote for low, and zero votes for insufficient. Therefore, the measure passes on validity.

Co-Chair Bratzler: Well, that was fast. Very good. All right. Starlin, we'll move to a discussion of feasibility, which again, just to remind everybody of the extent to which the specification including the major logic requires data that are readily available and can be captured without undue burden and implemented for its measurement. So Starlin?

Member Haydon-Greatting: So I'm proud to say that the great Medicaid turn to electronic data collection in 1990 after over '90 was passed has probably helped us with this measure because we collect more electronic information and claims are rejected earlier and sent back to get --

I just want to say that sometimes we have less missing data in these days because we can check -- we can check it faster but not as fast as you think because we're not in real time yet.

Anyways, for feasibility, I am confident that it's very feasible in Medicaid data. I'm concerned that it's a -- and I'm thinking that the commercial world would be able to feasibly apply this as well because they

probably collect more data than what the -- what the Medicaid government entities developed.

But I wanted to make sure I noted that all data elements needed to compute the measure are found in those defined bills.

There's even some more. There are G codes that you can pull to see if it's telehealth or telemedicine, and you probably should need to add that.

And that way, you can -- you can see what the follow-up was like and that would add some fidelity to the measure when you're -- when you're looking at that post.

There are some challenges in that linked claim data for emergency room. Again, I made the comment that sometimes emergency room hospitals make a deal with the Medicaid departments to get flat fee grants to cover some of what normally ambulatory care, primary care would cover in an underserved area.

So when I applied these, part of my master's thesis was on asthma in underserved areas of Chicago.

And when we applied these principles, we had to go and look individually at each hospital emergency room and see what the percentage they served as a primary care provider.

So that would have to happen in those geographically. So that may be an issue and you may be missing people because of that.

But all that are available, and they indicated in their notes that they're ready to operationalize this.

Now knowing that they also use Massachusetts and Vermont, that gives me a little more confidence about that.

And the staff rated this feasibility as moderate. No, high. No, moderate, sorry. Too many pieces of

paper.

Co-Chair Bratzler: Thank you, Starlin. Lindsay, do you have comments?

Member Botsford: I have the same concerns, just over identifying what claims would count, whether follow up visit, whether phone only, again, video or in person, may have some state variability in terms of the coverage of that as well.

So it could mean that different types of visits are included based on what states found as eligible claims based on their Medicaid criteria.

But all of that is gatherable through claims. So all of it is feasible. And no concerns with that rating of moderate.

Maybe just the comments to developers in a world of specifically value-based care arrangements, and wouldn't want to hamper creativity in ways of delivering care that meets the intent of this, which is to provide that education and follow up after an ED visit.

And wouldn't want to overly anchor to a certain type of visit to do so as an unintended consequence.

So maybe that's more on use than feasibility, depending on kind of what we include and what counts in those claims.

Co-Chair Bratzler: Thank you, Lindsay. Other committee members on feasibility? Naomi, just one quick question, because it's come up a couple of times. Have you tested this in other states?

Ms. Bardach:: We looked at it in the California, Massachusetts, and Vermont states, yes.

Co-Chair Bratzler: Okay. Any other comments from the committee?

Member Haydon-Greatting: So the benefit of doing

it in Massachusetts, they have a -- they are ahead, along with California, in ensuring their Medicaid delivery and electronic medical claim exchanges is higher.

Vermont probably is following along because they've tagged in with Massachusetts. What I wanted to say to the developers, the G code, because we did -- we did four diabetes during COVID.

We just updated all those CPT codes and ICD 10 codes and G codes so that there is a telephonic one, there is a telehealth one, and there is a telemedicine one, which means different things.

If it's a telemedicine, you're doing a remote monitoring with it, so there are all those codes you can pull. And asthma has some remote monitoring capabilities as well.

So those are things that you can pull in when you're doing your data claim pulls to see. Especially now in 2022 instead of in 2015. So I just wanted to add that. Sorry.

Co-Chair Bratzler: Thank you, Starlin.

Ms. Bardach:: Thank you.

Co-Chair Bratzler: Any other comments from the committee? Okay, let's go ahead with the vote on feasibility.

Mr. Wind: Voting is now open for Measure 3668 on feasibility. The options are A for high, B for moderate, C for low, and D for insufficient. Again, we will be waiting for 16 votes.

Voting is now closed. Measure 3668 on feasibility. We received three votes for high, 13 votes for moderate, zero votes for low, and zero votes for insufficient. The measure passes on feasibility. Thank you.

Co-Chair Bratzler: All right, thank you very much.

So Starlin, we're going to talk about use and usability next.

Use first, which is, is this metric being used by consumers, purchasers, providers, policymakers, for both accountability and performance improvement activities?

Member Haydon-Greatting: So right now it's not in public -- it's not in use across the board. They tested it in those three states with the Medicaid population.

My point is I think this is an excellent measure for the managed Medicaid. I often am frustrated with the quality of care managed Medicaid recipients are receiving.

So to enforce this with the managed Medicaid groups to make sure they have follow-up would be excellent.

It's kind of a political comment, but I review enough cases from the managed Medicaid group to know that they're lost in -- they're lost in follow up.

So again, they're ready to operationalize this and it probably can flow very well on Medicaid data.

I'm reserving commercial data, but if they pull enough of the right elements, they could create that pre-pull, the pull, and the post-pull, making sure all those ICD 9 codes and 10 codes are in there, because some people are still using 9s, and also the CPT codes that are purposely related along with G codes to see if anybody is receiving any hybrid post-follow up healthcare.

Co-Chair Bratzler: All right, thank you.

Member Haydon-Greatting: Lindsay, did you have anything?

Member Botsford: No, no concerns or extra comments on use.

Co-Chair Bratzler: So any other committee members on use? All right. Tristan, I think we can go ahead and vote on these.

Mr. Wind: Voting is now open for Measure 3668 on use. The options are A for Pass and B for No Pass. Again, we will be waiting for 16 votes.

Voting is now closed for Measure 3668 on use. It received 15 votes for Pass and one vote for No Pass. The measure passes for use. Back to you, Dale.

Co-Chair Bratzler: All right, thank you. Starlin, we'll go to usability, which I always struggle a bit with the differences, but it evaluates the extent to which audiences use or could use performance measure results for both accountability and performance improvement activities.

Member Haydon-Greatting: So the developer didn't share any potential harms or unintended circumstances.

I failed to see, if we're -- if we're encouraging follow up with the primary care, I guess the unintended consequence would be the person, the parents or the person with the disease, would be annoyed by having multiple calls, but that's not a bad circumstance.

That's keeping in touch with and getting engaged in your healthcare. So there's not much to go on.

I'm interested in knowing what other people's comments are on the usability in their professional practices.

Co-Chair Bratzler: Lindsay?

Member Botsford: Nothing, Dale. I mean, being a new measure, we don't have any data with regards to trends or how it has been used to improve performance. So the comments are speculative.

I think on the positive side, there positive

consequences would be encouraging health plans or other accountable entities to better at rapidly getting PCP or stakeholders notifications of ER utilization.

As pointed out earlier, this is tough to come by in the current world and often this information comes after the window has passed for gap closure.

So I think the incentives are right in terms of improving notifications to PCPs. Similarly, I think it could promote health plan outreach to ask-risk patients to encourage visits through care management or other disease-oriented programs. Again, positive.

Using unintended consequence probably doesn't outweigh the former, but unintended consequences actually with risk-bearing entities and value-based arrangements, health plan measures have a tendency to then be also used to evaluate PCP, especially performance as they're engaged in those relationships.

So I think there's a potential concern that the absence of great notification from a plan, and absence a lot of outreach, the onus would fall on PCPs.

I think especially in marginalized communities, as we pointed out, there was a disparity with lower follow ups by Black patients.

In current state, I think the risk could be PCPs who care for populations who have other barriers could follow up in current state, inability to take time off work, et cetera, I think the risk is that those PCPs would be evaluated as being poor performers on the measure.

Again, it's not a PCP, not a provider level measure in current state, but I think we have seen this with health plan measures, especially in certain types of relationships where it becomes a way that PCPs are

evaluations, which for things like this, benefits an affluent community.

So probably a fact within the Medicaid population normalizes the sample a little bit, although we all know it's still not perfect.

So those are the concerns. I'm not sure they outweigh the potential positive reviews. I do think they warrant discussion.

Co-Chair Bratzler: Okay, thank you, Lindsay. Other committee members? And, LeeAnn, I did get your notes and I will get to you here in a moment.

Ms. White: And, Dale, Kim Elliott has her hand raised.

Co-Chair Bratzler: Okay, thanks.

Member Elliott: Yes, I just wanted to say that was a Medicaid managed care rule that CMS has in place, the lack of requirements for notification of emergency department use will potentially not allow as much opportunity for useability of the measure because plans, the accountable entities, may not receive notice to be aware of that ease of use until past that 14-day timeframe.

So it just creates one little hurdle that makes it a bit more challenging.

Co-Chair Bratzler: Other committee members? So, Naomi, I'm going to -- and, then, LeeAnn, I'll get to you before we vote.

Naomi, I'm going to ask you to respond to that last comment.

Ms. Bardach:: I think it's an important point, how does this get operationalized to improve?

And as we discussed a little bit in the beginning part of the conversation, there's probably two pathways that can work, one of which is if somebody walks

into the emergency department, then they would get contact within, before that 14-day mark to be encouraged to get back into the primary care setting, either by the emergency department or by the health plan or by the PCP.

And having information flow is important. The other pathway to which this measure can drive improvement is actually just incentivizing health plans to help create those connections with primary care particularly for asthma patients.

And so, I think that other pathway, that would also be an important pathway for us to think about, is that that would be one part of the usability and absolutely something we would have to keep looking at as the measure gets more put into practice.

I'll say one other thing, which is that we don't have published, but Vermont was mentioning quality improvement work.

And they found that for their patients who were in a PCP office and then showed up in the emergency department later, it tended to be that they were patients who had uncontrolled asthma in the PCP office.

So there's a little bit of a pathway for PCP to say, oh, we have to pay a little bit more attention to those kids.

So that evidence base is still developing but that's another sort of mechanism through which this might help people focus a little bit more.

And I'll also ask Dr. Cabana, if you have anything else to add to the discussion.

MMember Curry: No, thanks, Dr. Bardach. Nothing else to add. Thanks.

Co-Chair Bratzler: All right. Thank you. Anyone else on the committee? Okay, so before we vote, I'm

going to ask LeeAnn to introduce Anna McCollister who has joined us. Anna, it's great to see you.

Ms. White: Thank you, Dale. Yes, Anna, welcome to the meeting. We just need to pause a moment before voting, just to have you introduce yourself.

Please, if you can come off mute, you can introduce yourself, your organization, and then any disclosures that you may have.

Member McCollister: Hi there. I'm Anna McCollister. I'm here as a patient advocate. I have no disclosures.

Ms. White: Wonderful. Thank you so much, Anna. Okay, I'll pass it over to, Dale are we ready for the vote?

Co-Chair Bratzler: I think we are ready for the vote.

Ms. White: Okay. I'll hand it over to Tristan.

Mr. Wind: Voting is now open for Measure 3668 on usability. The options are A for high, B for moderate, C for low, and D for insufficient. We will be needing 17 votes. We are waiting for two more votes.

Member McCollister: I'm having a hard time finding the link. Sorry.

Ms. White: Anna, we sent the link in the email around 8:30 a.m. Eastern Time. If you need us to resend that, we'd be happy to.

Mr. Sakyi: Anna, this is Isaac. You can also send your vote directly to me via chat.

Member McCollister: Okay.

Co-Chair Bratzler: All right, we're there.

Mr. Wind: All right, 17 votes. Voting is now closed for 3668 on usability. There was one vote for high, 15 votes for moderate, one vote for low, and zero

votes for insufficient. Therefore, the measure passes on usability.

Co-Chair Bratzler: All right. And so our last discussion is about overall rating of the measure for its overall suitability for endorsement. Is there any additional discussion at this point?

Member Haydon-Greatting: The only thing I'd like to bring up, that if you look at any of the competing -- NQF has 3599 Pediatric Asthma Emergency Department Use, and this measure is harmonized with it.

Co-Chair Bratzler: Okay. Yes, we will get to that. All right. Any other comments from any other committee members? Tristan, we'll turn it over to you.

Mr. Wind: Thank you, Dale. Voting is now open for Measure 3668 on overall suitability for endorsement.

The options are A for yes and B for no. And we will be waiting for 17 votes. One more vote.

Voting is now closed for Measure 3668 on overall suitability for endorsement. Therefore, the standing committee recommends to endorse the measure.

Co-Chair Bratzler: All right. Thank you. To our measure developers, thank you very much for being here to help answer our questions today.

We really appreciate it. And to the committee, I want to say you were amazingly efficient for our first measure of the morning. So good job, everybody.

Ms. Bardach:: Thank you very much. Take care.

Ms. White: Thank you so much. Yes, I echo, very efficient for our first one. So we have one more measure to review before our lunch break, and so I'm going to give a moment for Victoria to pull up

our slide and I'll introduce our next measure, which is 2797 Transcranial Doppler Ultrasonography Screening Among Children with Sickle Cell Anemia.

2797: Transcranial Doppler Ultrasonography Screening Among Children with Sickle Cell Anemia

The measure developer, Stewart Endovall, is from the University of Michigan. This is a maintenance measure and this is a measure that looks at the percentage of children ages 2-15 years old with sickle cell anemia who received at least one transcranial doppler screening within a year.

I will just actually pause a moment to see if our measure developer is on the line. So what I will do is I will hand this over to Adam, our co-chair, who will start leading the discussion.

We'll proceed like the last measure. And I will reach out to the developer to let them know that we are beginning their measure evaluation. So Adam?

Co-Chair Thompson: Thanks so much, LeeAnn. So thank you, everyone. We are going to begin to discuss this measure.

Our lead discussant is Mitch Harris and our supporting discussants are Kim Elliott and Brendle Glomb.

So, Mitch, would you like to begin us off with a brief introduction of the measure and then an overview of evidence?

Dr. Harris: Sure, thanks, Adam. So first thing, just to kind of mention, this was originally endorsed back in 2016.

So this is the first time that this measure is up for reindorsement. And I believe, again, it's a process measure.

It's again, looks at the health plan level. And as previously mentioned, looks at the percentage of

kids that are ages 2-15 that have sickle cell anemia and they received at least one screening with a year, a doppler screening within a year.

I don't think there's any sort of other information. I can just move right into evidence if you want, Adam.

Co-Chair Thompson: That'd be great. Thanks so much.

Dr. Harris: Sure. So again, a maintenance measure. And so we're just looking here to sort of see if there is any new evidence that has been provided.

In terms of the summary of the evidence that was in 2016, again, there was a systematic review of evidence for this measure.

There was some quality and consistency of the evidence that was provided. And the evidence was created based off guidelines from NHLBI, then again that supports screening among kids with sickle cell anemia.

There were, again, a couple of randomized trials and observational studies that were included in the evidence as well.

Again, back in 2016, the committee did note that there were some differences between age specifications in the guidelines but agreed that the measure aligned very closely with the NHLBI guidelines and sort of thought that that was okay.

There has not been any changes to the evidence since the measure was last evaluated. So the staff rated the evidence for the measure as moderate and the pre-evaluation comments, again, there were just some notes again stating that again, no changes to the evidence.

And some suggestions made that there's no need to repeat the discussion on evidence.

Co-Chair Thompson: Great. Thank you so much, Mitch. Kim, anything you would add?

Member Elliott: Everything I had reviewed, just, I was a little surprised that they hadn't done any updated evaluations or studies from 2016.

That's quite a bit of time to have additional evidence build up. But the evidence, they said there were no changes so I'm satisfied with that.

Co-Chair Thompson: Great. Thank you so much, Kim. Brendle, anything you would add?

Member Glomb: No, that was a concern of mine as well, with just the update of evidence and predictability, I suppose.

Co-Chair Thompson: Great. Thank you. Any questions or comments from the committee? All right.

Well, before we move to vote on the first area of evidence, we do have our measure developers that have joined us.

Julie, I believe you're on the line, if you'd like to give us a hello and welcome and a brief overview of your measure.

Ms. McCormick: Sure. Can you hear me? Are you able to hear me?

Co-Chair Thompson: Yes, ma'am. Yes.

Ms. McCormick: Okay. Terrific. I'm Julie McCormick. I'm the project manager at the University of Michigan.

Unfortunately, the measure developer, Dr. Sarah Reeves, is unable to join today. And I would like to note that I am going to have to leave at 11:30 and we sincerely apologize for our limited availability today.

But I would like to give a brief statement about the measure. As noted, the measure assesses the percentage of children ages 2 up to 16 years old with sickle cell anemia who received at least one transcranial doppler screening within a year.

This measure is supported by strong evidence. Without intervention, 11 percent of children with sickle cell anemia will have a stroke by the age of 18.

Importantly, these strokes can be prevented. Transcranial Doppler, or TCD screening, is a non-invasive ultrasound method to identify children who are at high risk of stroke by measuring the blood vesicles and the vessels of the brain.

Among those of higher risk, receipt of chronic blood transfusions dramatically reduces the risk of stroke, a 92 percent reduction was observed in a randomized control trial.

Given the fact that TCD is the only method with which to identify children with sickle cell anemia at the highest risk of stroke, an expert panel at the National Heart, Lung, and Blood Institute, or NHLBI, strongly recommends that all children with sickle cell anemia should receive one TCD per year from ages 2 up to the age of 16.

Therefore, in concordance with the NHLBI guidelines for annual TCD screening, our measure uses administrative claims to assess the proportion of children ages 2 up to 16 years with sickle cell anemia that receive a TCD screening within a year.

As you know, there are many hemoglobin variations. Our measure focus on HBSS cases, sickle cell anemia, which is consistent with NHLBI recommendations.

Our denominator is the number of children with sickle cell anemia. We have tested the case definition for this denominator in both ICD 9 and

ICD 10 by comparing various case definitions to the gold standard of newborn screening and choosing the most accurate definition.

Therefore, this case definition is valid as it has both sensitivity and specificity of approximately 90 percent to identify children with sickle cell anemia as compared to newborn screening record.

Our numerator is the number of children with sickle cell anemia that received a TCD screening.

Receipt of TCD screen is identified through the presence of an administrative claim for the test.

Again, we found that this method was valid, as it is highly correlated with TCD screens found in the medical record.

In addition to the strong evidence and validity of this measure, its application identifies significant opportunities for improvement in care among this vulnerable population.

Since 2010, only about 40 percent of the children with sickle cell anemia received TCD screens when this measure was applied to Michigan and New York state's Medicaid programs.

This measure has shown no increase through 2019. The measure is currently in use in quality improvement initiatives, particularly in the state of Michigan.

Working with the University of Michigan, the state Medicaid program has established a pilot program, a collaborative of Medicaid programs within a specific region of the state, designed specifically to improve the care of children with sickle cell disease.

As part of this collaborative, Medicaid health plans within the region will receive financial incentives from the state if performance rates for this measure improve over the course of the next year.

Finally, it is important to acknowledge that the majority of children with sickle cell anemia are within underrepresented and underserved racial and ethnic groups in the U.S.

This measure directly speaks to health disparities experienced by this population as compared to other chronic conditions.

In closing, we feel that this measure is highly important and has the potential to make a substantial, positive impact on the quality of children's lives.

Our measure focuses on the receipt of TCD screening, which is the only method in which to identify children with sickle cell anemia that are at highest risk for stroke.

We found that this measure is highly valid and that the data elements to calculate this measure are readily available and administratively straightforward.

We found that there is an important performance gap, and this measure is currently in use to incentivize Medical health plans in Michigan to improve the quality of care for children with sickle cell anemia.

We strongly believe that endorsement of this measure will have a very positive impact on the health of these high-risk children. Thank you for your time and consideration.

Co-Chair Thompson: Thank you so much, Julie. And Kim?

Member Elliott: When you say that it's currently in use in one region in Michigan Medicaid program, what percentage of the Medicaid population then is included in that region? Do you have that data?

Ms. McCormick: I don't have the exact number, but it is Region 10, which is Southeast Michigan, where

a majority of the children with sickle cell anemia receive their care.

Co-Chair Thompson: Great. Thank you so much. Any other questions or comments from the committee? Okay.

And if no one has any objection to the evidence, because this is a maintenance measure, there's no change, we can move directly to gap, unless somebody has an objection.

If anybody has an objection and would like us to vote on evidence, you are welcome to let us know that.

And you can either send a note to LeeAnn and Isaac and let them know or let us know here in the chat room.

LeeAnn or Isaac, have you received any objections?

Ms. White: I have not.

Mr. Sakyi: I haven't.

Co-Chair Thompson: All right. Well, then, I think we can go ahead and move onto the opportunity for improvement. Mitch?

Dr. Harris: Sure. Thanks, Adam. In terms of gap again, no new information provided in the new paperwork.

But again, looking back at information from 2016, there were rates provided from six states from both 2005 and 2010 that showed some slight improvement in these states, again, from a low, some places 5 percent and maybe up to 51 percent over the five years.

Again, and then there's also some information that looked at gaps in two Medicaid programs, one in Michigan and one in New York that also showed sort of, even though there was some improvement,

there still was gap that remained with rates below 50 percent.

In terms of information that was on any potential disparities, again, they did not have any data in terms of gender, income, or socio-economic status.

But they did sort of look at information by age and it was clear in some of the younger ages they had a higher screening rate compared to children that were in the older ranges.

In terms of some of the pre-evaluation comments, again, there was just that sort of a gap and opportunity of improvement still exists, or at least in the data previously existed.

And the staff indicated the opportunity for improvement currently at moderate.

Co-Chair Thompson: Great, thank you so much, Mitch. Kim, anything you would add?

Member Elliott: No, I concur with everything that Mitch has said. There is opportunity.

The level of performance is still relatively low. And all of the words, socioeconomic, age, gender, any of those sort of disparity breakdowns, there's clearly still evidence that supports opportunity for improvement.

Co-Chair Thompson: Thanks so much, Kim. And Brendle?

Member Glomb: I do not. I think the demographic is set on this. I do think that this is a room for improvement for managed care across the board.

And given the population, I think a lot of them are represented within the Medicaid population.

So I think that this is, it's actually something I'm going to take back to work and make sure that we're pursuing actively through the breakdown and

very useful measure. It's a little long in the tooth.

Co-Chair Thompson: Great. Thank you so much, Brendle. Any other questions or comments from the committee members?

All right. Hearing none, I think we can move to the vote on gap. Tristan?

Mr. Wind: Thank you, Adam. Voting is now open for Measure 2797 in performance gap. The options are A for high, B for moderate, C for low, and D for insufficient. We will be requiring 17 votes. Waiting for one more. Perfect.

Voting is now closed for Measure 2797. There were five votes for high, 12 votes for moderate, zero votes for low, and zero votes for insufficient. Therefore, the measure passes on performance gap.

Co-Chair Thompson: Great. Thank you so much. Moving onto scientific acceptability, beginning with reliability. Mitch?

Dr. Harris: Thanks. So this measure was not reviewed by the Scientific Methods Panel. So any comments that provided again were provided by staff previously.

The numerator and denominator statements appear straightforward. Again, there are no denominator exclusions.

And so again, as the measure developer indicated, the numerator is children age 2-15 with sickle cell anemia who received one TCD screening within the measurement year.

So again, it does look within that measurement year. And then the denominator is just the number of children in that age range with sickle cell anemia.

In terms, and again, so no denominator exclusion. Again, in terms of specification, they're really straightforward and I don't think there are any

issues brought up by the evaluation comment.

In terms of the reliability testing, again, this was done at the accountability entity level.

They used the CMS max data from about 2005 to 2012 for six states, and again, looked at some reliability statistics. And they were in the high 90 percent range. So again, the developer indicated there was a high degree of reliability.

Again, committee stated that the specifications, they believe they can be consistent and had no concerns.

And again, the committee also indicated that it looks reliable and so there may not be a need to repeat the discussion and vote on reliability.

The preliminary rating for reliability from the staff is moderate.

Co-Chair Thompson: Great. Thank you so much, Mitch. And Kim, anything you would like to add?

Member Elliott: No, I just wanted to comment that since they were using the CMS max data, why it would only tie back to the accountable entity versus the state Medicaid program or health plan level, because I think that would have been easy.

Well, maybe not easy, but a way that you could have looked at the data as well going past the 2016 time period.

So I would encourage the developer to maybe broaden that a little bit so that there's a higher degree of paper performance or other sort of methodology that could be put in place if it's more broadly used.

Co-Chair Thompson: Thanks so much, Kim. And Brendle?

Member Glomb: I have maybe a follow up question.

Have we learned where in the system the failure to improve the goal exists?

Co-Chair Thompson: Any responses from our other discussants before we ask the measure developer at the end of this session?

All right. Any other questions or comments, and then, Brendle, we'll bring your question over to the measure developer?

Member Glomb: Thanks so much. Where's the breakdown, right?

Co-Chair Thompson: All right. Seeing there are no other questions or comments, Julie, any response to Brendle's question?

Dr. Parmac: Yes, actually, we have an entire program. Working on that to answer that question right now.

We're seeing a lot of difficulty in being able to reach the families of children with sickle cell anemia in Region 10.

Right now we see there is a 60 percent, approximately 60 percent unable to reach rate for these families.

So we're working with case managers and community health workers within the region on outreach.

We're also, if children aren't being contacted, then we don't know if they're seeing their physician, their primary care physicians, or a hematologist. So that's part of the work we're doing now.

Co-Chair Thompson: Great. Thank you so much, Julie. Brendle, any follow up? All right. Any other questions or comments from the committee?

All right. So seeing how this is a maintenance measure, we can choose, LeeAnn, keep me honest

here, we can choose to accept the previous committee's decision that this measure is in fact reliable and move forward to validity unless someone has objections, which you are free to let us know now or message to Isaac or LeeAnn.

All right. Isaac, LeeAnn, we good? LeeAnn, thumbs up. Let me scroll down. I'll let Isaac interrupt me if he gets anything.

Let's go ahead and move forward then onto validity. Mitch?

Dr. Harris: So again, just note that there has been no new validity testing updated for this measure.

So again, all these information provided was from the original sort of endorsement back in 2016.

There were three levels of validity testing conducted at that time, and there was a patient encounter level that looked at chat reviews and looked at the ability for the reviewers to agree and that sort of had some high scores.

Then there was also sort of validity testing done at the accountability entity level, using some of the Michigan max data and then some of the Michigan Medicaid claims and data and then comparing those and they seem to have a high level of reliability between the two sources.

And then there was an expert panel, again, that looked at face validity of this measure and rated it very highly and thought it would improve the care of patients with sickle cell anemia.

So again, no new information provided at this time. In terms of the pre-evaluation comments, no significant concerns were brought up with the validity testing or any other potential threats to validity, including the lack of risk adjustment or the lack of exclusions in this measure.

The preliminary rating for this validity is moderate.

I'll turn it back over to you, Adam.

Co-Chair Thompson: Thanks so much, Mitch. Anything you would add, Kim?

Member Elliott: No, it's a rather straightforward measure. No risk adjustment or stratification.

They didn't identify any missing data elements. It was claims driven. And I think that the measure, the way it's designed, can distinguish between poor and good quality care, so no concerns.

Co-Chair Thompson: Thanks so much, Kim. And Brendle, anything you would add?

Member Glomb: I do not. Thanks. Thank you, Adam.

Co-Chair Thompson: Thanks so much. Committee, any questions or comments for our discussants on the measure?

Dr. Harris: Adam, this is Mitch again. And just maybe sort of one comment to the developer.

Given their statement last time about sort of the trouble reaching out and being able to contact some of the children with or the families of children with sickle cell anemia, I'm wondering why they don't have any risk adjustment based on some socioeconomic data.

There are some new metrics out there like the Child Opportunity Index or some others that look at some area-based things that maybe would help if there are true differences in terms of some of the perhaps issues with being able to contact families perhaps that live in specific areas or maybe that have resource limitation. So that may be something to look at in the future.

Co-Chair Thompson: Thanks so much, Mitch. Any other questions or comments from the committee? All right.

So we did have updated validity information so we do need to take a vote on this criteria.

So I think, no comments or questions at this point, we can move to validity testing vote. Tristan?

Mr. Wind: Thank you, Adam. Voting is now open for Measure 2797 on validity. The options are A for high, B for moderate, C for low, and D for insufficient. We will be waiting for 17 votes.

Voting is now closed for Measure 2797 on validity. There was one vote for high, 16 votes for moderate, zero votes for low, and zero votes for insufficient. Therefore, the measure passes on validity.

Co-Chair Thompson: Thank you so much, Tristan. And moving next to feasibility. Mitch?

Dr. Harris: So this measure seems pretty feasible to put into place. All data elements are defined fields and electronic claims data.

Didn't seem to be any concerns from the pre-evaluation comments. Again, indicating that the data was available in electronic claims and seemed available and very accessible.

And then the pre-rating was moderate by the staff.

Co-Chair Thompson: Thanks so much. Kim, anything you would add?

Member Elliott: No, I agree. The fields are accessible. The codes exist. It's all electronic claims or claims, so no concerns.

Co-Chair Thompson: Thank you. And Brendle?

Member Elliott: Again, I think this is easily performed.

Co-Chair Thompson: Great. Thanks so much. Great, thanks so much, Brendle. You were breaking up there a little bit but I think it sounds like you got a

little bit better there at the end.

Any questions or comments from the committee for our discussants on the measure? Any questions folks have on feasibility for the developer?

All right, I think we can move to a vote on feasibility. Tristan?

Mr. Wind: Thank you, Adam. Voting is now open for Measure 2797 on feasibility. Options are A for high, B for moderate, C for low, and D for insufficient. Again, we will need 17 votes.

Voting is now closed for Measure 2797 on feasibility. There were six votes for high, 11 votes for moderate, zero votes for low, and zero votes for insufficient.

Therefore, the measure passes on feasibility.

Co-Chair Thompson: Great. Thank you so much. And now moving to use and usability, beginning with use. Mitch?

Dr. Harris: Thanks. So I think this is, we're getting into the exciting part of this measure now. So a couple of issues that will come up in the next two.

I think in terms of use, again, previously it's been identified that it is not being used in public reporting.

It is being used in accountability programs and that there is planned use for accountability programs as well.

Again, it's being used within the Michigan Medicaid program for a payment program to improve rates.

And while not being currently publicly reported, the developer indicated that there are plans to be used in a quality improvement program, which I believe they mentioned earlier, to try to increase rates as well.

The developer does provide the results to Michigan Medicaid Health on a quarterly basis. So again, there is -- there is certainly some use. It is in accountability, but not currently within sort of publicly, being publicly reported.

Pre-evaluation comments focused on, again, the identification that it's being used in accountability, not public reporting, but no sort of other significant issues.

And the preliminary reporting for this measure for this staff was a pass.

Co-Chair Thompson: Great. Thanks so much, Mitch. And Kim?

Member Elliott: I would just like to put a question out there for the end of the discussion, Adam, on which other publicly reported program the developer is working with. Otherwise, no concerns.

Co-Chair Thompson: Thanks so much. I'll put that question on the list, Kim. And Brendle?

Member Glomb: Out here in Texas, it's not a part of public reporting. I am wondering about accountability, though, and to my comment a while ago about this, I'm going to follow up internally here and see where we are both with accountability from state and feds as well as how much we are holding our providers accountable.

Co-Chair Thompson: Great. Thanks so much, Brendle. Before we ask our measure developer Kim's question, any questions or comments from the committee for our discussants?

All, right hearing no -- Vilma?

Dr. Joseph: Yes, again, I was just curious about that six-year window. So we are at that mark in terms of public reporting. Did it have any barriers to it at all? And they can talk about it.

Co-Chair Thompson: Any response from our discussants before we move to the measure developer?

All right, any other questions or comments for the committee before we turn with our two questions for the measure developer?

All right. Julie, are you still with us?

Dr. Parmac: Yes. I will need to leave shortly, but I'm happy to do my best to answer your questions.

Co-Chair Thompson: Awesome. So I think we've got two questions sitting here on the table.

One of them is about the six-year window that Vilma was just talking about, and the other is Kim's question related to, I've gotten written down and it's scribbled and I can't even read it.

Member Elliott: Which other accountability --

Co-Chair Thompson: There you go, which other accountability programs?

Member Elliott: Right.

Co-Chair Thompson: Thanks so much.

Dr. Parmac: Sure. Well, this is probably a question, both questions are probably best answered by Dr. Reeves, but I will do my best.

I understand that if this pilot program in the Southeast Michigan is successful in increasing rates, it is something that will be rolled out.

There have been discussions about rolling out through the state of Michigan. So that's one issue.

There's also discussion being made at the national level to possibly have this measure included in the core set of measures, CMS core set of measures.

Those discussions have been ongoing for a few

years now. It has been recommended three times for inclusion in the core set, although that has not happened. So that is something that is still an ongoing discussion at the national level.

Co-Chair Thompson: Thanks so much. And Vilma, did you get your answer in there? All right, well, we will circle back around if we lost Vilma's question.

All right, any other questions or comments from the committee about use? All right, Tristan, I think we can move to the vote.

Mr. Wind: Thank you, Adam. Voting is now open for Measure 2797 on use. Options are A for Pass and B for No Pass. We will be needing 17 votes.

Voting is now closed for Measure 2797 on use. There were 17 votes for Pass and zero votes for No Pass. Therefore, the measure passes on use.

Co-Chair Thompson: Thanks so much, Tristan. And usability, Mitch?

Dr. Harris: So this is the area where I think there might be some work to do. But I'd love to hear from the developer, too, based on sort of what she said about the improvement project and if they maybe have some preliminary data that they could share.

And again, while it is being used for different accountability programs, not, I'm sorry, the program, it's not really -- information on improvement has not been shared over time.

So really no information to be able to judge the usability at this point in time during the maintenance phase.

So pre-evaluation comments, again, indicated that there certainly aren't any unintended consequences or harms because of the measure, but again the lack of performance data calls into question the usability at this point.

The preliminary rating from staff again on this measure was insufficient.

Co-Chair Thompson: Thanks so much, Mitch. Kim, anything you would add?

Member Elliott: Yes, I agree with everything that Mitch just said. I am questioning why more information hasn't been put out in the last six years since this measure was first approved or recommended, endorsed by NQF.

And if it's been in used in Michigan in the one region, Region 10, is there any data or information available as to whether the quality improvement work that's being done and the measurement resulting from that is resulting in improvement in care or outcomes for this individual.

Co-Chair Thompson: Great. Thanks so much, Kim. And Brendle?

Member Glomb: I agree with Kim completely. I think we're past time for some more updated data for substantiation of the measure.

I support the measure, but it does seem like we've got to, at some point we've got to be able to be sure that the measure is relevant and is being used as it should be.

Co-Chair Thompson: Thanks so much, Brendle. Any questions or comments from the committee for our discussants?

Dr. Harris: Is she still here? I know she had to leave at 11:30.

Co-Chair Thompson: I was just looking to check my participant list. That was my little secret. I don't see her on here anymore. Julie, are you still with us?

Dr. Parmac: Yes, I'm still here.

Co-Chair Thompson: Awesome. Do you want to

speak as much as much time as you can give us right now about sort of the lack of updated performance improvement data?

Dr. Parmac: So we do have data that we did not have available to share with you at the time that we made the submission, but we do have six quarters of data for the state of Michigan, specifically in Region 10 where the rates for this measure continue to be around 36 percent. Those last five or six quarters include January through March of 2022.

We've initiated this collaborative with the state of Michigan. We are working diligently to change these rates and have these improvement programs in place.

We do expect to see results over the next year. The state of Michigan will be starting the incentive program for July 2022 to July 2023.

We have all of these reports in place to hopefully make changes in the care and in the performance rates for this measure.

Co-Chair Thompson: Thanks so much, Julie. Any follow up questions or comments from the committee?

Co-Chair Bratzler: I guess my only comment is that when I reviewed the evidence, and it's not a field that I follow closely, I don't take care of kids, but I was impressed by the potential risk for stroke that needs to be addressed.

And honestly, there are a number of sickle cell metrics that we'll mention later in the related, that are, I just don't know that anybody's holding anybody accountable for them.

They seem like they're really important. I can't tell you how many kids we have in the emergency room all the time with sickle cell anemia.

So I would hope that people would look at these

carefully. If everybody feels strongly that the evidence is good, these should potentially be used for accountability measures.

Co-Chair Thompson: Yes, thanks so much, Dale. Any other comments or questions from the committee? All right, I think we can move to the vote, Tristan.

Mr. Wind: Voting is now open for Measure 2797 on usability. The options are A for high, B for moderate, C for low, and D for insufficient. Again, we will need 17 votes. Waiting for one more vote.

And voting is closed. Voting is now closed for Measure 2797 on usability. We received five votes for high, seven votes for moderate, one vote for low, and three votes for insufficient. Just give the team a moment to confirm the votes.

The measure passes on usability. Back to you, Adam.

Co-Chair Thompson: Thanks so much. So next up is overall suitability for endorsement before we discuss related and competing measures.

So if there's any final comments or questions the committee has before we move to that final vote --

Dr. Harris: Yes, this is Mitch. Again, I'll just say that I think that as we've gone through, it's a certainly a measure that seems like there's a lot of evidence for but it seems fairly simple and straightforward to apply.

But as Dale just mentioned, I think the challenge is understanding why there's not more of a pickup in the use of this.

So it may be sort of just find leaders to different organizations or groups who can implement it.

I do again think on the earlier points that we voted on, there's certainly a lot of evidence for this

measure to be in place and more accountability in public reporting systems.

Co-Chair Thompson: Thanks so much, Mitch. Any other comments from the committee? I think we can move to overall suitability for endorsement.

Mr. Wind: Thank you, Adam. Voting is now open for Measure 2797 on overall suitability for endorsement. Options are A for yes and B for no, and 17 votes.

Voting is now closed for Measure 2797 on overall suitability for endorsement. There were 17 votes for yes and zero votes for no. Therefore, the standing committee recommends to endorse the measure.

Co-Chair Thompson: Thanks so much, Tristan, and I think now, a brief discussion on our related or competing measures, beginning with Mitch.

Dr. Harris: Sure. So there were a couple of related measures indicated. Again, 3166, Antibiotic Prophylactics in Kids with Sickle Cell Anemia, and 3595, Use in Children with Sickle Cell Anemia.

The developers indicate that they are harmonized with the current measure to the extent possible.

They are looking at the pre-evaluation comments from the committee, again, people agreed with that assessment and believes that there was pretty good harmonization.

Co-Chair Thompson: Great. Thanks so much, Mitch. And I've been informed that I am way too excited about this, to discuss related and competing measures, which we will discuss at the very end of our meeting because between now and that part is actually your lunch break, so I don't want to be that guy that steals your lunch break for a conversation we're going to have later.

So we'll put a pin in that when we come back to it. Thank you so much, Mitch, for providing that

preview of our discussion that we will be having a little bit later. And I will pass it back to LeeAnn.

Ms. White: Adam, I love your enthusiasm to keep rolling. So, yes, I will -- while we give everyone the lunch break, let's rest our eyes for a little bit, step away from the computer, get a good stretch in, get some food, we'll reconvene at 12:30 and we will start with, we have two composites.

So definitely a good break and then we'll come back. We'll do 0729 and then 3294. And we'll wrap up the day.

So, I appreciate everyone's participation. It is going super smooth and efficient.

The voting is going great. So thank you all for being so engaged this morning. And I look forward to seeing you after lunch. All right. Thank you.

Ms. White: Okay. It is 12:31 p.m. on the East Coast, so we will go ahead and get started. Welcome back, everyone, to our Spring 2022 PCCI Measure Evaluation Meeting. Glad to see you all back.

We will move into our last two measures of the day. And so, I'm going to pause and allow Victoria to share her screen so we can move to our third measure, which is 0729. Thank you, Victoria.

0729: Optimal Diabetes Care

So, our next measure is 0729, optimal diabetes care. The measure steward and developer is Minnesota Community Measurement. This is a composite measure.

The measure is the percentage of patients 18 to 75 years of age who have had a diagnosis of Type 1 or Type 2 diabetes, and whose diabetes was optimally managed during the measurement period, as defined by using all of the following: hemoglobin A1C less than 8.0; blood pressure less than 140

over 90; on statin medication unless allowed contraindications or exceptions are present; non-tobacco user; and patient with ischemic or vascular disease on daily aspirin or anti-platelets, unless allowed contraindications or exceptions are present.

So, with that, I will go hand it over to our Co-Chair Adam Thompson to lead us in the discussion.

Co-Chair Thompson: Thanks so much, LeeAnn. As we begin, I believe we have our measure developers with us, Collette and Julie. If you would like to give us a brief overview of the measure before we begin our discussion.

Dr. Cole: Great. Thank you. Good afternoon, everyone. I'm Collette Cole, a measure developer with Minnesota Community Measurement. And with me is Julie Sonier, our president of MNCM. We're pleased to be presenting Measure No. 0729, Optimal Diabetes Care for consideration of the endorsement.

This measure is a patient-level, all or none composite measure that seeks to reduce modifiable risk factors associated with the long-term macrovascular and microvascular complications of both stages of diabetes.

Patients with diabetes are more likely to reduce their overall risk, prevent or reduce complications, and optimize health outcomes by simultaneously achieving several intermediate physiological targets and medication adherence conformance.

LeeAnn did a great job of going over the measure, the denominator and its components. So, I'm going to skip that part.

These components are treated equally. There is no weighting of individual components.

The scientific acceptability was reviewed and approved by both the Scientific Methods Panel and

the PCCI Committee during the 2018 indoor cycle.

The measure has not changed since the last endorsement review. Therefore, scientific testing was not updated for this submission.

In its preliminary analysis, NQF staff indicated in discussion during the 2018 maintenance review regarding the evidence supporting the use of all five components to improve outcomes. We'd like to address any potential concerns with the following information:

The American Diabetes Association standards of care in 2022 have an A level recommendation for the optimization of glucose and blood pressure control to reduce the risk or slow the progression of chronic kidney disease.

In 2018, a study of over 270,000 patients with Type 2 diabetes in Switzerland found that the excess risk of cardiovascular illness outcomes decreased in a stepwise fashion for each risk factor variable that was within target range.

Patients who were successful in achieving targets for all five risk factors have little or no excess risk of death, myocardial infarction, or stroke as compared to the general population.

And, lastly, Health Partners, a large, integrated health system in Minnesota, has reduced the incidence of long-term complications of diabetes. Risk per thousand members between 2000 and 2017 fell from 17.8 to 9.1 for acute myocardial infarctions; from 4.8 to 4.3 for amputations; and from 68 to 40.3 for retinopathy.

For Health Partners and their approximately 42,000 members with diabetes, annually this equates to 361 fewer MIs, 20 fewer amputations, and 954 patients not experiencing retinopathy.

This measure, originally developed by Health

Partners, has been publicly reported in Minnesota for almost 20 years, and is included in the Minnesota Department of Health statewide quality reporting and measurement system, which requires patient-level data submission for all primary care and endocrinology clinics in Minnesota.

On an annual basis, we are receiving data on over 315,000 patients with diabetes in Minnesota and from other communities.

In 2020, the statewide average was 40.6 percent of all patients achieving all five targets, increased from our initial rate of 9.5 percent in 2006. Potential effects of the pandemic were noted, with a rate decrease of 4.8 percent between 2019 and 2020.

There is significant variability among practices with the lowest scoring clinic at 5 percent, and highest scoring clinic at 56.1 percent.

Additional variability is highlighted when stratifying the measure by race, gender, ethnicity, insurance type, and neighborhood socioeconomic variables.

Groups that have the biggest rate percent decreases were female, those age 40 to 49, the uninsured, and higher socioeconomic status.

The measure still demonstrates a gap in opportunity to improve.

Please note, that while the measure is an all or none composite. The individual components may be measured as well, and are particularly helpful in quality improvement efforts to better understand where opportunities for improvement exist.

Thank you for the opportunity to present this measure for your consideration.

Co-Chair Thompson: Thank you so much for that overview. We really appreciate you all coming today.

Let's go ahead and turn to our committee discussions. Just a reminder, because this is a composite measure we will have a couple other criteria that we need to discuss.

Let's begin with importance to measure specifically evidence. And our lead discussant is James, and then we'll follow up with Ann and Robert.

Member Rosenzweig: Hello. Can you hear me?

Co-Chair Thompson: Yes, we can hear you.

Member Rosenzweig: So, this measure is a longstanding measure. It's been used since 2006. And it's been used specifically in the State of Minnesota.

And it has been associated, at least we don't know if there's any causative effect of the use of the measure in improving these outcomes, but we've seen the substantial improvement of the measure itself over time, and correlating with individual components of them.

Do you want me to go through the actual components again?

Co-Chair Thompson: I don't think we necessarily need to do this at that point, unless committee members have a specific question. Because it is a maintenance measure, when we get to evidence we can, as a committee, just accept the evidence vote that we had in our previous endorsement.

So, if folks have particular questions about it, if you want to put that in the chat room. If not, I think we can kind of go ahead and keep going with a higher level review.

Member Rosenzweig: Okay. The one specific thing that was listed in the measures is that they -- in the material you sent me it says hemoglobin A1C less than 8.0 milligrams per deciliter. And that's a mistake. It should be 8.5 -- 8.0 percent. It's a

percent. The hemoglobin A1C is measured as a percentage, not as milligrams per deciliter.

Okay. So, this composite has two intermediate outcome measures and three performance measures.

Just to note again to the committee, 20 years ago I was the chair of the Operations Group of the National Diabetes Quality Improvement Alliance which developed the first nationally-recognized performance measures for diabetes care. We existed until 2009, and disbanded then.

But this measure actually uses our materials back then in 2006. The thing about it is that a lot of the various measures that have been developed since then have changed, whereas this has been pretty constant since 2006.

Because I have no conflict of interest because it was so long ago.

And this is an all or none composite measure. The plan is that this should be considered the gold standard, reflecting best patient outcomes. However, of course, individual components might be measured as well. And then made with similar individual measures for patients with diabetes related to microvascular outcomes.

The developers don't adequately explain why a composite like this should have specific benefit. They refer in the staff, and they refer to a Swedish study from 2018 that cited they reviewed almost 300,000 patients with diabetes overtime, showing stepwise improvement in macrovascular complications in each of their targets as measures. However, the components are different from those in the measure under consideration.

Their targets in that study were blood pressure of 148 over 80, not 140 over 90, which is in our measure.

The hemoglobin A1C cutoff was 7.0 percent as opposed to 8.0 percent.

Smoking, and then they used smoking rather than tobacco use.

Their LDL and the cholesterol was used, so they had LDL cholesterol of greater than 97. Whereas we're just -- our measure says whether or not they're taking a statin, period. No reference to specifically the degree of hyperlipidemia.

And then, in addition, they used a presence of albuminuria as their fifth component, rather than our using the use of aspirin as an anticoagulant.

Their outcomes are risk of death, myocardial infarction, stroke, and heart failure. So, they really focused on the macrovascular complications instead of the microvascular complications.

However, we all know from numerous studies, like a DCCT, KPDS, and EDIC that glycemic control is clearly associated with improvement of vascular and microvascular complications as well. And blood pressure has a big input.

So, basically they are just we see comparisons of a composite measure achievement in 2006 to two thousand -- which improved substantially from 2006 to 2016.

Then they stayed steady until 2020. And then they significantly declined in 2021, likely due to COVID-19.

Now, the evidence they cite are the evidence of -- they also, they also show, I can mention now, the clear differences in performance by race, ethnicity, socioeconomic state, insurance, as well as other factors.

Now, the evidence that they cite are the ratings in the 2022 ADA standards of medical care for diabetes. However, the ADA recommendations do

not include a specific A1C target anymore. They had originally.

And they give a general goal of 7 percent and not 8 percent, with individual goals to be determined for each patient, either higher or lower than 7 percent based upon clinical factors, like increased risk of hypoglycemia, limited life expectancy, et cetera.

And they also encourage the use of shared decision making regarding targets, which actually is summarized in the staff review document.

The ADA statin recommendations are more nuanced than the ones in this measure, but they are generally reflective.

The low dose aspirin, the low dose aspirin is recommended by ADA for those with existing ASCVD or increased risk for ASCVD. This measure substitutes increased cardiovascular risk with just one item, which is LDL cholesterol greater than 190.

And the blood pressure control recommendation in the ADA recommendations they indicate 140 over 90, only for those patients with low risk for the presence of ASCVD. Patients with existing CVD or 10-year ASCVD risk greater than 15 percent should be treated with a more highly -- with a lower target of 130 over 80 safely obtained.

So, this is a composite measure, but it clearly does not indicate optimal diabetes care. When it was, perhaps back in the day when it started out, it did. But it would make sense to call it more comprehensive diabetes care or adequate diabetes care.

And it is true that if you're looking from year to year, usually these, you know, these measures, whenever the cutoff is, they tend to, they tend to go along with the, you know, they tend to track along with other, other cutoffs.

Now, I just also want to mention there are not many important measures of diabetes care that are not included here. They include eye exams at least every two years and more frequently, measurement of kidney function and macro, micro or macro albuminuria.

There's no reference to any weights, BMI or weight circumference, since weight loss is clearly a very important target of Type 2 diabetes care.

Then there is no mention of any behavioral interventions. And no measures of diabetic dyslipidemia. Patients with high LDL -- excuse me, low HDL and high non=HDL cholesterol, or high triglycerides are very important factors as well.

So, the best information we can glean from the measure is the long-term information it provides, since it was originated in 2006.

And because they didn't change the measures over time, if they had, we wouldn't be able to compare performance from year to year.

And for this reason, it should be considered a legacy measure. It makes sense to retain this measure for those who use it currently, but there's a big question as to whether or not it should be used for new programs starting out.

Do you want me to discuss performance gap, opportunity for improvement, or should I stop here?

Co-Chair Thompson: Yeah, no, we can stop at evidence because it's a different vote before we move forward.

Thank you so much, James.

Ann, anything you would add?

Member Kearns: No. I don't, I don't think so. I agree that it's not really anymore considered optimal care. There are components left out.

Like, I don't think they talk about peripheral nerve evaluation, or foot exam, or things that we all know lead to complications or amputations.

So, you know, but it is trying to get at a more comprehensive low bar probably.

So, I don't have anything to add to that.

Co-Chair Thompson: Great. Thanks so much, Ann.

And before we turn to the committee, Robert, anything you would add?

Member Bailey: I'm aligned with everything that's said so far.

But, on the other hand, even though it's not optimal, there's still a significant opportunity to improve performance against this measure that addresses the two major areas of morbidity and mortality of the patient population. Right? So, microvascular disease and macrovascular disease.

Co-Chair Thompson: Great. Thanks so much.

So, turning to the committee now, just before we start making comments, when we get to the end of this I see, Anna, that you've got your hand up, so I'll come to you first.

We can accept the evidence that was presented earlier as part of our vote. So, just keep that in the back of your mind as we've having this discussion. Or we can bring the evidence back up and take a full committee vote.

Anna.

Member McCollister: Hi there.

I have never been comfortable with this measure since it was first proposed to us back on the Endocrine Standard Committee when I was a member of that, which preceded this committee.

One of the issues that I have with it, although I have to say, I mean, I have Type 1 diabetes. I have taken statins. Have taken, you know, 10 milligrams of Lipitor for more than 20 years at this point. I take low dose aspirin, have for more than 20 years.

You know, my A1C is in that range. Blood pressure. So, I understand the benefits and I understand the rationale for these, for each of these components very well.

However, as the daughter of two people with Type 2 diabetes, one of whom has had a history of severe reactions, like, very severe reactions to statins. Because of quality measures that have been drummed in that has every PCP and cardiologist throughout the country, they keep putting her back on statins.

A couple years go by to see her reactions, they just put her back on statins. Nobody ever does any pharmacogenetic testing. I've had her tested. She does have a reaction to statins. Nobody takes it into consideration unless you, like, throw that over their heads.

There's never any discussion of her -- anyway, if we're going to have these measures for a drug for which we know that 30 percent of the population has a tendency towards a severe adverse event, we have to have this into the measure, very specific challenger measures that require and also measure the ability of a physician to be able to think through pharmacogenetic testing. Or at least, if not PGS, which is readily available, asking about side effects and tracking side effects very closely.

Because it's really easy for a physician to think, oh, well, the patient has a bad outcome for a couple of months or three months, so we take the drug, so we stop the drug and she goes on something else. That's the name of life that are very substantial, can be problematic, and can interfere with compliance with other medical regimens. They generally

complicate with age and with life.

And, I mean, these measures for how long at this point, and just with all of the issues that were pointed out previously, there's still a significant failure to address this issue. And given out we needed to address that issue, the fact that it still hasn't been taken into consideration in the context of this measure just really I find to be problematic.

Member Rosenzweig: I'd just like to thank Anna for all the work she's done for us over the years. I mean, I remember her commenting on this 10 or 15 years ago.

And the thing about the issue with statins is that, is that since this measure started there are a lot of other medications that are available for treatment of LDL cholesterol. Some of the data is not as clear cut as with statins in terms of reducing macrovascular events. But, still, now we usually put someone on another medication if they're intolerant to statins. And nothing is specifically discussed on that regarding this measure.

Co-Chair Thompson: Thank you all.

I did want to draw attention to the chat room. Our measure developer Collette put in there that the measure has an exception for allergy and intolerance to statins, and aspirin, and antiplatelets.

So, I just wanted to put that there.

And then, Kim, I saw your hand go down. Does that mean you were going to point to the chat room, too?

Member Elliott: Yeah, that's exactly what I was going to point out, that they do have that exception methodology for people with issues with statins, et cetera. Yep.

Member McCollister: But there's no guidance to, like, have a physician do pharmacogenetic testing or

specifically test for or inquire about adverse events. It just has that exception.

And, you know, the denominator which, great, also there's no conventions or requirement to, like, follow up on, you know, any of those issues or to do preventative testing. It just incentivizes people to use a drug or alternative.

Co-Chair Thompson: Thanks so much, Anna,.

Any other questions or comments from the committee for our discussion related to evidence?

(No response.)

Co-Chair Thompson: Okay, hearing none, so we have two options, folks.

One, we can accept the previous vote of the committee around evidence and move into importance to measure in the gap.

Or, we can reopen the evidence vote and vote again.

If anyone has objections to continuing forward to the next measure and accepting the previous vote, if you could let us know. Either send LeeAnn or Isaac a note, or let us know in the chat room if you would like to vote on evidence. Let me make that clear, because I know that was a lot.

If you're okay moving forward and not voting on evidence, let us know.

Member Rosenzweig: One thing I should measure is that -- one thing I should mention is that they are using a different, they are using the ADA guidelines now and evidence guidance for that. Whereas, prior to this they were using another set of conform -- set of diabetes conformance measures that basically were derived from the ADA.

So, I don't know the differences between the older

set, but that other organization, and I think it's a Minnesota organization, that other organization is no longer in existence.

Co-Chair Thompson: Thanks so much, James.

And we did get some folks who would like to vote on evidence.

So, before we move to the vote, any last comments or questions from the committee on evidence?

(No response.)

Co-Chair Thompson: Okay. Tristan, can we go ahead and bring up the vote.

Mr. Wind: Thank you, Adam.

Voting is now open for Measure 0729 on evidence. The options are A for moderate, B for low, and C for insufficient.

We are looking for 17 votes.

(Pause.)

Mr. Wind: One last call for votes.

Ms. White: Is anyone having difficulties casting their vote? If so, you can directly message myself or Isaac.

Mr. Wind: And voting is now closed for Measure 0729 on evidence.

There were 13 votes for moderate, 2 votes for low, and 1 vote for insufficient.

Please provide us one moment.

(Pause.)

Mr. Wind: The measure passes on evidence.

Thank you, Adam.

Co-Chair Thompson: Thanks so much, Tristan.

Moving forward, importance to measure looking at opportunity for improvement for gap.

James.

Member Rosenzweig: Okay. So, there are a number of performance gaps that are clearly demonstrated from the information they gave us.

Statewide results show that 55 percent of patients diagnosed with diabetes have at least one component of the measure that was not optimally managed. And I discussed the Swedish study.

But all five components, people who actually had all five components of this measure had extremely low risk of major complications in their study.

The main component for mortality is smoking in other studies, whereas here we're talking about tobacco use which, I guess, includes all sorts of other things, like vaping and things of that sort.

Regarding the impact, I think we're seeing very interesting data regarding the impact of COVID-19. Statewide measure decreases from 45 percent in 2019 to 40.6 percent. But more than 5 -- more than 10 percent decline.

And these rates declined significantly among females, a loss of 4.8 percentage points. The age 40 to 49 group, 5.4 percentage points. The uninsured, 12.9 percentage points, a tremendous decline. And the high, and people with high socioeconomic status had a fairly -- had more than a lot of other people, they were actually 6 percent point decline since the beginning of the pandemic.

And the rate changes ranged from only a slight decrease of each individual component, only ranged from a very slight decline of 0.1 percentage point for tobacco, but a 7 percent decline overall for blood pressure control and hemoglobin A1C results, very

significantly increased in 2020.

With regard to disparities, the developer also provided data related to COVID. Asians declined 6.3 percent; Blacks declined 5.4 percent; Whites declined 4.6 percent; and indigenous natives 2.7.

Non-Hispanic declined 4.8 percent, whereas Hispanic declined less, 4.4 percent.

So, there's a wide range of performances among different ethnic groups. And I think this data is interesting and useful for follow-up.

Okay. I think that -- is that the quality, sufficient for discussion of the quality construct?

Co-Chair Thompson: No. We'll move to that part next.

Member Rosenzweig: Okay. Thank you.

Co-Chair Thompson: Yeah, thanks so much, James.

And before I turn to Ann, I just want to note, in the chat room Collette, from the measure developer, did note that tobacco is referring to tobacco products only and does not include vaping products. So, they don't consider those tobacco use. So, just to clarify there.

Member Rosenzweig: Okay.

Co-Chair Thompson: Ann, anything you would add?

Member Kearns: No. I think that's a nice summary. And it is impressive, the impact of, presumably, COVID.

Co-Chair Thompson: Thanks so much, Ann.

Robert.

Member Bailey: Just to emphasize. There is still significant opportunity for improvement here.

Co-Chair Thompson: Great. Thanks so much.

Any questions or comments from the committee for our discussions about opportunity for improvement?

(No response.)

Co-Chair Thompson: All right. Hearing no questions or comments, Tristan, I think we can move to the vote.

Mr. Wind: Thank you, Adam.

Voting is now open for Measure 0729 on performance gap. Options are A for high; B for moderate; C for low; or D for insufficient.

(Pause.)

Mr. Wind: We're at 15 votes.

Last call. And if you have any technical difficulties, please contact LeeAnn or Isaac.

Last call.

Voting is now closed for Measure 0729 on performance gap.

There were 6 votes for high; 10 votes for moderate; 0 votes for low; and 0 votes for insufficient. Therefore, the measure passes on performance gap.

Co-Chair Thompson: Great. Thank you so much, Tristan.

And, again, because this is a composite measure we do have a third criteria in this section, and that is looking at quality construction and rationale.

James.

Member Rosenzweig: Okay. The idea, of course, is that a combination of measures can indicate a broader attention to overall care rather than individual measures considered separately.

Each of the five components of the measure is important in and of itself. But the combination is, indeed, more effective at reducing complications. However, two of the measures are restricted to sub-population.

So, aspirin use really applies only to those with existing risk factors of complications for those between 20 years of age and 40, whereas over 40 it's for all of the people with existing complications.

And then there is no weighting of the components of the construct. That can be a significant issue, you know. There's no -- because some of the components are more targeted towards specific negative outcomes.

And the most recent hemoglobin A1C in the measuring period is less than 8.0. It applies to all denominator patients.

And the most recent blood pressure, less than 140 over 90, applies to all the denominator patients.

The statin use, if appropriate, and no complications, exceptions.

Diabetic age -- diabetics with age 18 to 20 have a free pass. And those with existing ischemic vascular disease on statin, unless LDL is less than 40, are exceptions in that category.

And then diabetic age, 21 to 39, and LDL greater than or equal to 190, that's for patients between 21 and 39.

So, and then people who have an LDL less than 190, which is the vast majority, they'll get a free pass.

So, so they're different, so it's kind of a little bit problematic in that each of the measures don't exactly measure the same populations. They're measuring different populations. Okay.

And the daily aspirin is also a complicated measure. It's only for those patients with diabetes who actually have ischemic vascular disease.

We used to actually recommend daily aspirin for most people with diabetes. But now it's clear that a lot of the risks of daily aspirin, especially in young people, actually outweigh the benefit. So, so it's been cut back to an older age population.

So, the general grade on this I would give would be moderate.

Co-Chair Thompson: Great. Thanks so much, James.

Ann, anything to add?

Member Kearns: I think he summed up the issue quite well.

Co-Chair Thompson: And Robert?

Member Bailey: Nothing else to add, thanks.

Co-Chair Thompson: Great.

And to the committee, any questions or comments from folks for our discussion?

Member McCollister: I would just say moderate at this point is being generous. I mean, measures will live with us for a long time. We have had science come out in the intervening years. I think we need to take this, like the new understanding of aspirin's potential risks versus lack of benefits for aspirin into consideration.

I think we need to take even a more highly developed understanding of statins under consideration.

And I don't think it's adequate to just say, you know, put a thing in the denominator to say that if people can't take those medications or if the risk

factors aren't there, whatever the situation may be, then those people are not included in the numerator, because that stuff gets overlooked if a doctor is just trying to adhere to quality measures to make the numbers.

So, we have to take responsibility for the unintended consequences on these measures as well as the potential benefits. And this is just no longer -- I think it's been tenuous from the beginning and I just don't think it (audio interference), so.

Co-Chair Thompson: Great. Thanks very much, Anna.

And I think, Collette, I think you had your hand up.

Dr. Cole: Yes. Hi. This is Collette, Minnesota Community Measurement.

And if I may, I'd just like to address some of the comments.

So, yes, this measure has been in place for quite some time. However, we constantly are looking at the new evidence as it's arising. And we have convened three different work groups to review that evidence.

For example, when the measure started out, the A1C component was less than 7. With the results of Accord and DCCP, we rapidly convened a work group and looked at all of the recommendations and possibilities, and understanding that trying to capture some events like limited life expectancy, hyperglycemia, and other elements that would keep some patients from achieving a less than 7 target, our work group decided to go with less than 8.

Subsequently, when the guidelines changed for cholesterol management we had a more targeted physiological target of an LDL less than 100. But with new study, that was no longer supported. And

that was substituted with the use of statins. And based on recommendations from the ACCAHA.

And, likewise, when new blood pressure recommendations came up, I want to say about 2016, we again convened the development work group and reviewed that evidence and, you know, came to the determinations that we have today.

In terms of the couple of components, for example the aspirin, that changed in evidence and guidelines. Previously, the recommend -- or the target was diabetic patients age 40 and older were on daily aspirin. However, that with more evidence is not a safe thing.

So, we modified that component to just look at patients with known cardiovascular events, and looking for aspirin on that. But in terms of the measure construct, patients who don't have ischemic vascular disease, they are -- they are not dinged or pinged for not taking aspirin, because they shouldn't be. So they're given a free pass on that component.

So, I'd be happy to answer any other questions about kind of the evolution of this measure over time.

Thank you.

Co-Chair Thompson: Thanks so much, Collette.

Any other questions or comments from the committee related to quality construct?

(No response.)

Mr. Wind: Hearing none, I think we can move to the vote, Tristan.

Mr. Wind: Thank you, Adam.

Voting is now open for Measure 0729 on composite quality construct and rationale. Options are A for

high; B for moderate; C for low; and D for insufficient.

(Pause.)

Mr. Wind: We are at 16 votes. Last call for a vote.

Last call.

Voting is now closed for Measure 0729 on composite quality construct and rationale.

There were 0 votes for high; 15 votes for moderate; 1 vote for low; and 0 votes for insufficient. Therefore, the measure passes on composite quality construct and rationale.

Co-Chair Thompson: Thanks so much.

Moving on now to our next criteria, scientific acceptability. This will again have three sections to is: reliability, validity, and then empirical analysis.

So, let's begin with reliability. And first, James.

Member Rosenzweig: Okay. So, there were no changes to the specifications of the measure since the last review. And I think that's appropriate because it's very important not to change things in midstream when you're undergoing, you know, when you have a situation like COVID.

And the reliability testing level was both on accountability, entry -- entity level, and also on the patient encounter level.

So, reliability correlated with a number of eligible patients in the clinic, and ranged from 0.51 -- 519 in the clinics with the maximum patients of 30 -- minimum patient of 30. And to the larger clinic it was 0.994.

So, they indicate a overall high level of reliability for the measure score. So, I'd rate this high.

Co-Chair Thompson: Thanks so much, James.

And, Ann, anything to add?

Member Kearns: No.

Co-Chair Thompson: No.

And Robert?

Member Bailey: Nothing else to add. Thanks.

Co-Chair Thompson: Thanks so much.

Any questions or comments from the committee for our discussions on reliability?

(No response.)

Co-Chair Thompson: All right. Seeing none, I think we can move to the vote.

Mr. Wind: Voting is now open for Measure 0729 on reliability. Options are A for high; B for moderate; C for low; and D for insufficient.

(Pause.)

Mr. Wind: We are at 16 votes. Last call for a vote.

Last call.

Voting is now closed for Measure 0729 on reliability.

There were 6 votes for high; 9 votes for moderate; 0 votes for low; and 1 vote for insufficient. Therefore, the measure passes on reliability.

Co-Chair Thompson: Thanks so much.

And moving now to validity. James.

Member Rosenzweig: Okay. So, validity testing at the patient encounter level, they validated, the developer validated the elements by performing an audit and quality check of 53 medical groups; 30 percent of those submitting data.

And a total of 89 percent of the medical groups

passed the initial audit. 11 percent required a correction plan. And of those medical groups who submitted data, all passed the audit with greater than 90 percent accuracy.

On the accountability -- on the accounted -- Excuse me.

On the accountability entity level, I guess that means, like, health plans and so forth, the developer conducted validity testing for a composite score. And by testing the correlation of medical group performance with their performance of the overall measure.

And so, and they hypothesized that the, they hypothesized that the quality of care would be, should be similar. Then analysis of the medical groups' performance on the optimal diabetes care measure demonstrated a fairly strong correlation with its performance on optimal vascular care with respect to the measure itself.

So, it was like a correlation coefficient of 0.629.

Then they also conducted validity testing for the individual components of the composite and using Pearson R correlation analysis. For hemoglobin A1C the correlation was 0.78. For blood pressure the correlation was 0.71. For tobacco free was 0.54. For statin use, 0.68. But aspirin or anti-platelet use was only 0.26254.

So, that's pretty much what I have here for you.

Co-Chair Thompson: Thanks so much, James.

Anything you would add, Ann? No.

And Robert?

Member Bailey: Nothing else, Adam. Thanks.

Co-Chair Thompson: Thanks so much.

Any questions or comments from the committee for our discussion?

(No response.)

Co-Chair Thompson: All right. Seeing none, Tristan, I think we can move to the vote.

Mr. Wind: Voting is now open for Measure 0729 on validity. Options are A for high; B for moderate; C for low; and D for insufficient.

(Pause.)

Mr. Wind: We are at 15 votes. Last call for voting.

Last call.

Voting is now closed for Measure 0729 on validity.

There were 2 votes for hi; 12 votes for moderate; 1 vote for low; and 1 vote for insufficient. Therefore, the measure passes son validity.

Co-Chair Thompson: Great. Thank you so much, Tristan.

Moving to our last section here around empirical analysis. James.

Member Rosenzweig: Empirical analysis.

The component, the individual component measures add value to the composite. And that's aggregating and weighting rules consistent with the quality construct.

This is an all or none composite measure. Each component -- I sort of mentioned this before -- each component reduces the modifiable risks associated with diabetes.

The desired goal of the composite measures is for a patient to achieve intermediate physiological outcomes and medication use targets to best decrease their overall risk of developing

microvascular and macrovascular complications related to diabetes. The developer used Pearson product moment correlation to measure the strength of linear regression of the relationships between the composite and its components.

I think I mentioned this earlier actually. But, yeah, so there was strong correlation with four out of the five components: blood pressure, hemoglobin A1C, and statin use, and tobacco use. But aspirin use was very low.

There were 618 clinics with 306,000 patients correlation of performance -- No. I don't need to do that.

Co-Chair Thompson: Hold it. We're getting to there in just a second. Thank you, James.

Ann, anything you would add?

Member Kearns: No.

Co-Chair Thompson: No.

Robert?

Member Bailey: Nothing else, thanks.

Co-Chair Thompson: All right. Thank you all.

Andy questions or comments from the committee related to empirical analysis.

(No response.)

Co-Chair Thompson: All right. I think we can move to the vote, Tristan.

Mr. Wind: Thank you, Adam.

Voting is now open for Measure 0729 on empirical analysis. Options are A for high; B for moderate; C for low; and D for insufficient.

(Pause.)

Mr. Wind: Fifteen votes. Last call for voting.

Last call.

Voting is now closed for Measure 0729 on empirical analysis.

There was 1 vote for high; 12 votes for moderate; 1 vote for low; and 1 vote for insufficient. Therefore, the measure passes on empirical analysis.

Co-Chair Thompson: Thank you so much.

Moving to our next criterion, feasibility. James.

Member Rosenzweig: Okay. The measure has been used in Minnesota without problems since 2006. The data is accessible and easily obtained.

The number of patients and clinics are adequate.

And the developer is implementing a new electronic data warehouse system. And there are no fees associated with the participation for data submission.

So, I would say feasibility is high.

Co-Chair Thompson: Thanks so much, James.

Anything you would add, Ann? Nothing.

And Robert?

Member Bailey: I'd just add that the data is captured in the routine delivery of care, so it doesn't provide any additional burden to clinicians.

Co-Chair Thompson: All right. Thank you so much.

Any other questions or comments from the committee for our discussion on feasibility?

(No response.)

Co-Chair Thompson: All right. Hearing none, Tristan, we can move to the vote.

Mr. Wind: Voting is now open for Measure 0729 on feasibility. The options are A for high; B for moderate; C for low; and D for insufficient.

(Pause.)

Mr. Wind: Sixteen votes. Last call for voting.

Voting is now closed for Measure 0729 on feasibility.

There were 12 votes for high; 3 votes for moderate; 0 votes for low; and 1 vote for insufficient. Therefore, the measure passes on feasibility.

Co-Chair Thompson: Thank you.

Moving now to use and usability, beginning with use.

James.

Member Rosenzweig: Use. It's actively used in appropriate locations in the state.

Co-Chair Thompson: Awesome.

Ann, anything you would add?

Member Kearns: I agree.

Co-Chair Thompson: And Robert?

Member Bailey: Agree.

Co-Chair Thompson: Any questions or comments from the committee for our discussions on use?

(No response.)

Co-Chair Thompson: All right. We can move to the vote.

Mr. Wind: Voting is now open for Measure 0729 on use.

Options are A for pass; B for no pass.

(Pause.)

Mr. Wind: Sixteen votes. Last call for voting.

Voting is now closed for Measure 0729 on use.

There were 15 votes for pass; and 1 vote for no pass. Therefore, the measure passes on use.

Co-Chair Thompson: Thank you.

And moving now to usability. James.

Member Rosenzweig: Okay. The developer basically the measure is included in two state regulatory programs, the Minnesota Statewide Quality and Reporting System, and it's also included in the Minnesota Health Care Home Certification/Recertification Program.

So, the developer provides statewide performance gap data for achieving all the five components of the composite measure, as well as the individual performance rates in each of the five components.

So, statewide rates increased from 9.5 percent in 2006 to 53.5 percent in 2015 which is a, you know, very substantial improvement. But then a decline -- it remained level until about 2019, and then it went down.

So, it's now 45.4 percent, or at least in 2020 it was.

And then there was a decrease in most of the various outcomes, which I measured before, various weights of the components.

So, benefits versus harm, is that included in this as well? So, the benefits of the performance measure in facilitating high quality, efficient care outweigh the incidents of unintended negative consequences to individuals or populations.

The unexpected findings, there were two unexpected findings. Adults age 65 and older with

Medicare have better outcome rates than younger adults with diabetes due to generational differences related to providers' orders.

And then statewide A1C averages are trending upward, which is a trend that the ADA has confirmed.

And, also, of course, I gave information earlier about the impact of COVID-19. And this is important because COVID-19 and diabetes are linked in the sense that people with diabetes have worse, tend to have worse outcomes when they are infected with COVID-19.

And in certain situations COVID-19 can actually aggravate problems with diabetes control.

Co-Chair Thompson: Great. Thanks so much, James.

Anything you would add, Ann? No.

Robert?

Member Bailey: Nothing else, thanks.

Co-Chair Thompson: Any other questions or comments from the committee for our discussions on usability?

(No response.)

Co-Chair Thompson: All right. I did just want to myself bring back on the comment, I think, what Anna was talking about earlier, the potential harm associated with prescribing the statins for folks who have the reaction to it.

I know it wasn't mentioned in the measure developer, but I just want to kind of bring that forward again in this section.

All right, if we have no other questions or comments, I think we can go ahead and move to

the vote.

Mr. Wind: Thank you, Adam. Voting is now open for Measure 0729 on usability. Options are A for high; B for moderate; C for low; and D for insufficient.

(Pause.)

Mr. Wind: Fifteen votes. Last call for voting.

Last call. Voting is now closed for Measure 0729 on usability. There were 5 votes for high; 10 votes for moderate; 0 votes for low; and 1 vote for insufficient. Therefore, the measure passes on usability.

Co-Chair Thompson: Great. Thank you so much. Last section, overall suitability for endorsement.

Any comments or questions from the committee members around the measure in its totality?

Member Rosenzweig: I'd like to say something.

Just, you know, overall there are a lot of problems with this measure. And but I think it should be continued, certainly in this population.

The question is whether or not we can recommend, NQF can recommend it for use or endorse it for use elsewhere. I would be unlikely to want that to happen. So, I'm not sure we can make that kind of exception. So, it's just a point I would like to make.

Member McCollister: What do you mean by that NQF can't endorse it for use elsewhere?

Member Rosenzweig: Well, I think that a new composite measure should have different, would be better if it had different criteria within the composite. The five categories are not necessarily the best five categories to look at.

And but, you know, and also the cutoffs are different from what we would necessarily like as

optimal diabetes care.

Nevertheless, it's been very useful in Minnesota. And I think especially during, during the period of COVID now that we're following, we really need to see follow-up data for the next few years to see where things are going.

Co-Chair Thompson: Thanks so much, James.

Member McCollister: I'll just reiterate what I said before. I mean, I don't know how many times I need to say it. Like, the consequences on individual outliers can be significant. And I don't, I have never felt like a measure developer, or NQF, or community has taken those consequences as seriously as they should be.

It's super easy from the perspective of public health to overlook the consequences on individuals when they're outliers. It's just really easy.

And I don't know, I'm not, like, a measure developer expert, I don't have any kind of secret sauce which in terms of, like, dealing with outliers. But when you put together composite measures that have no compensatory requirements or any kind of risk assessment or mitigation baked into the measure, any position dealing with an elderly patient, like, some level of lack of respect and dismissiveness that gets translated in that environment is significant. Even if the patient is educated, articulate, they get ignored.

So, anyway, I just need to evolve our approach to how we define and measure quality. And I feel like we're not doing that. And it's very frustrating.

Co-Chair Thompson: Thanks so much for that, Anna.

Any other comments from the committee?

Ms. White: I just want to let Anna know that we definitely appreciate your feedback on the measures

and the review. And when we do review the measures as specified, we can definitely take these recommendations and your feedback and note those in our report for the developer.

We just want to clarify that we can't endorse for specific use. So, just to clarify that. But we can -- but we, the standing committee endorses measures as specified, and those recommendations can be noted. And we will make sure we note those in our summary and our draft report.

So, thank you for that.

Co-Chair Thompson: Thanks so much, LeeAnn.

Any other questions or comments?

(No response.)

Co-Chair Thompson: All right. I think we can move to the vote for overall suitability.

Mr. Wind: Thank you, Adam.

Voting is now open for Measure 0729 on overall suitability for endorsement.

Options are A for yes; and B for no.

(Pause.)

Mr. Wind: Fifteen votes. Last call for voting. Last call.

Voting is now closed for Measure 0729 on overall suitability for endorsement.

There were 13 votes for yes; and 2 votes for no. Therefore, the standing committee recommends to endorse the measure.

Thank you, Adam.

Co-Chair Thompson: Thanks so much. And we will be voting on related and competing measures, no

matter how excited I get about them, at the end of our process.

So, I think now I'm going to be passing it back to LeeAnn.

Ms. White: Thank you, Adam. So, we are about to mark on our last measure under review for the spring cycle. So, I'm going to -- while Victoria pulls up our slides, we see our measure here.

3294: STS Lobectomy for Lung Cancer Composite Score

So, this measure is a surgery composite measure, Measure 3294: STS Lobectomy for Lung Cancer Composite Score. The measure steward and developer is the Society of Thoracic Surgeons, This is a maintenance measure.

I will provide a brief description of the measure. The STS lobectomy composite score comprises two domains: operative mortality during the same hospitalization as surgery, or within 30 days of the procedure. Presence of at least one of these major complications: pneumonia, acute respiratory distress syndrome, bronchopleural fistula, pulmonary embolus, initial ventilator support greater than 48 hours, pre-intubation respiratory failure, tracheotomy, myocardial infarction, or unexpected return to the operating room.

The composite score is created by a weighted combination of the above two domains resulting in a single composite score. In addition to receiving a numeric score, our systems are applying to rating categories designated by star ratings, one star, two star, and three stars.

So, I am going to hand the baton over now to our Co-Chair Dale Bratzler to help lead us through the discussion of Measure 3294.

Co-Chair Bratzler: All right, finally. I'm having

mouse problems; I couldn't get the mute button pushed. So, thank you for that introduction. Not sure who we have on from the developer, so maybe Dr. Michael Firstenberg will give the brief introduction of the measure to start with. If Michael is here, or one of the other members of STS?

Ms. White: So, Michael is our lead discussant, but I do believe we have members of our STS developer.

Dr. Yagci: Yes, LeeAnn. Hi, this is Banu Yagci with the Society of Thoracic Surgeons. Here with me today I have Dr. Jeff Jacobs, Dr. Dave Shahian, and Dr. Moritz Wyler von Ballmoos, we're very happy to be here, and after Dale, Dr. Jacobs will proceed with the introduction with our measure.

Dr. Jacobs: Hi. So this is Jeff Jacobs, it's nice to be talking to you all today, and thank you for your service in the National Quality Forum. I think the previous introduction to the measure said the majority of what I had plans to say anyway, but what I can tell you is that this is measure maintenance on a multi domain composite.

It's one of many multi-domain composite quality measures in the STS portfolio of measures, and this is one that deals specifically with lobectomy for lung cancer, which is the most common surgical procedure done in patients with lung cancer. This composite measure has been reviewed, and published in the peer review literature, it was published in the Annals of Thoracic Surgery in 2016.

And the measure was developed by utilizing an analysis of the STS general thoracic surgical database, as was stated before this is a composite score that is developed based on two outcomes, risk adjusted mortality, and any, or none risk adjusted presence, or absence of major complications. The model was developed utilizing data from the general thoracic surgery database over a four year analytic window.

And 95 percent Bayesian credible intervals were utilized to determine, and categorize performance with a star rating system utilizing three stars. The analysis was performed with a study population of 20657 patients undergoing lobectomy for lung cancer at 231 hospitals across the United States. In that analytic cohort, the operative mortality was one, and a half percent, and the rate of major complications was 9.6 percent.

The median post-operative length of stay was four days. Risk adjusted mortality, and major complication weights varied three fold from the highest performing three star providers to the lowest performing one star providers. And after placing all this information into the star system, the multi domain composite categorized hospitals as 88 percent in the two star domain, five percent in the one star domain, and seven percent in the three star domain.

After publication of this analysis in the Annals of Thoracic Surgery, this measure was incorporated into the STS general thoracic surgery database, initially it was utilized for feedback to participants in our feedback reports that are distributed every six months to the participants, and subsequently was also included in the portfolio of measures that the Society of Thoracic Surgeons publicly reports.

So, I think that's a quick introduction of this measure, and there's three STS surgeons, as well as Banu from the STS office on the phone, and we're all happy to answer any questions that may arise during the discussion, thank you.

Co-Chair Bratzler: All right, thank you Dr. Jacobs, I appreciate that. So, I'm sorry, the lead discussant today is Dr. Michael Firstenberg, so Michael, are you on the call?

Dr. Firstenberg: Yes.

Co-Chair Bratzler: Okay, great, I'll turn it over to

you to discuss the evidence.

Dr. Firstenberg: I mean it's great to see everybody again, and Jeff, it's good to see you again, seems like we're kind of tag teaming these every couple months.

Co-Chair Bratzler: It's good to see you Michael.

Dr. Firstenberg: Too, I guess to kind of be brief out of respect for everybody's time, this is a pretty well validated model within the STS that runs parallel to many of the cardiac metrics that are publicly reported, and tracked in a similar fashion. The data is kind of collected from all the participant sites while it is somewhat voluntary, it is sort of encouraged that everybody participates in this procedure submits all their data.

It's been validated, it's been used for years, and it's hard for me to really improve on the particular presentation that Dr. Jacobs put together, but this is probably as robust as they get for a quality metric for the surgical treatment of early stage lung cancer. I think it incorporates a very robust statistical modeling, highly effective data collection that's been validated again, many times over the years.

And is publicly reported, published in the peer literature, and is something that is pretty well respected by everybody at this point. Probably it in a nutshell.

Co-Chair Bratzler: All right, thank you. Our supporting discussant, Miklos Kertai.

Dr. Kertai: Thank you, I concur with what Dr. Firstenberg said, it's robust, well validated, proven to be very useful risk prediction tool. So, I have no concerns.

Co-Chair Bratzler: Okay, Dr. D'Agostino?

Dr. D'Agostino: I echo the comments, and I have no

concerns either.

Co-Chair Bratzler: All right, thank you. Anyone else on the committee have any questions, or comments about the evidence for the metric? All right, hearing none, I believe we can move to voting on the evidence.

Mr. Wind: Thank you Dale. Voting is now open for Measure 3294 on evidence. The options are A for pass, and B for no pass. We are at 15 votes, last call for voting. Voting is now closed for Measure 3294 on evidence. There were 15 votes for pass, and zero votes for no pass. Therefore the measure passes on evidence.

Co-Chair Bratzler: All right, thank you. So, Michael I know in the introduction, Dr. Jacobs talked a little bit about performance gap, if you would highlight performance gap.

Dr. Firstenberg: Well, I think this is something that has been validated over the years. I think one of the challenges, the variability in the number of cases that are presented from programs range from kind of small, intermediate, and large, but I think that again, that has been pretty well developed within the star rating program, that is consistent.

And so, again, it kind of captures all of the information, and I don't think there's any substantial performance gap in any of this.

Co-Chair Bratzler: I think maybe in the performance gap --

Dr. Firstenberg: I'm sorry?

Member Rosenzweig: There should be a performance gap.

Dr. Firstenberg: Yeah, so there are some concerns about just the -- sometimes the under representation according to this, of the demographics in terms of who gets presented,

particularly those with different types of insurance, whether it's Medicaid, or private insurance, but I don't think that this is anything that warrants -- the question is whether there's a national performance measure.

Whether there's any opportunity for improvement, I think we just need to continue to capture the data, and I think it would be the disparity related to the number of cases, does sometimes potentially impact smaller programs that may do less volume, and that statistically one, or two cases may throw off their numbers.

But I think that is somewhat adjusted in the risk adjustment models that are presented. Does that address I think what you're looking for?

Co-Chair Bratzler: I think what we're thinking about when we talk about performance gap, is there still an opportunity to improve performance on the measure? And I see Dr. D'Agostino shaking his head yes, and I believe that's what I heard from Dr. Jacobs.

Dr. Kertai: Yeah, that is correct, I believe there is an opportunity to improve here, the developer, as it is indicated, and the evaluation document identified a significant disparity between bias, and non-bias in terms of the performance, and outcome. So, indeed there is a performance gap which provides an opportunity for improvement of caring, and providing care for minorities.

Even more fundamental than that though, I would say that until we have all patients with lobectomy having no complications, and no deaths, there is a performance gap, and we are substantially nowhere close to that.

Dr. Jacobs: Yeah, that's exactly what I was going to say. I think the data shows that there's still an important number of hospitals that perform at the one star level on this measure, either in one

domain, or the other domain, or in the overall composite, and that represents a performance gap with opportunities for improvement.

Dr. D'Agostino: Yes, I agree with that entirely. The other thing to point out is the performance gap that you highlight here in terms of the ethnic backgrounds, I think it's also important to understand that the thoracic surgeons are downstream in the referral process. So, they will get their patients referred to them from pulmonologists, and primary care, and internists.

So, some of the gaps that you see in terms of the ethnic backgrounds are reflective of the care, and the referral practices of the referring physicians. But clearly there's an opportunity for improvement, and an opportunity for outreach to our referrings.

Dr. Firstenberg: I think the other comment that was brought up, and maybe others can address this better, is the fact that this really just captures the majority of the data I think performed by thoracic surgeons. And I think that there may be a lot of these types of operations, or similar, or lesser operations that get performed throughout the country by general surgeons.

And how much of that data is captured, and how much of that is accurately reflected in the overall data set.

Member Glomb: I'm going to check, and see if there are other members of the committee that have questions, or comments. I'm going through some of the pre-evaluation comments. Other members of the committee who have comments, or? Okay, well I think we can move forward with voting on performance gap.

Mr. Wind: Voting is now open. For Measure 3294 on performance gap. Options are A for high, B for moderate, C for low, and D for insufficient. 15 votes, last call for voting. Voting is now closed for

Measure 3294 on performance gap. There were five votes for high, ten votes for moderate, zero votes for low, and zero votes for insufficient. Therefore the measure passes on performance gap.

Co-Chair Bratzler: All right, we'll move forward. Michael, discussion of the quality construct with focus on the composite itself?

Dr. Firstenberg: Yeah. So, the composite construct looks at not just operative mortality during either the indexed hospitalization, or within 30 days, but a variety of complications that are pretty well described, and defined. And those then get used to risk adjust everything into the different levels of star ratings.

With regards to this, the question then is, including some of the current data as to the rule of minimal base of lung surgery, and whether that can continue to improve overall outcomes in morbidity, and mortality, and whether additional measures can be included to try to optimize this.

Co-Chair Bratzler: Dr. Kertai?

Dr. Kertai: I agree with Michael. And in addition to that, I'd like to add that as we discussed, and briefly described by Dr. Jacobs, this is a robust quality construct, and there is proven evidence over the years that there is basically -- how the rational, how the risk prediction model was designed, and validated.

And there's also, at the time basically publications, and all of the above indicated that this is a clinically useful prediction model, which can withstand the test of time.

Co-Chair Bratzler: Okay, Dr. D'Agostino?

Dr. D'Agostino: Yes, I agree. I think this is a very robust model, and includes the not only mortality, which obviously is the most significant thing, but

also encompasses the very important complications that are meaningful, and have been shown to impact outcomes. So, this is -- I have no concerns on this at all.

Co-Chair Bratzler: Any questions from the committee on the construct of the composite?

Dr. Joseph: This is Vilma, I like the fact that they changed it over time. Before it was one to one mortality versus complications, and now it's four to one mortality versus complications, and so I thought that was pretty insightful, that they wrote in the change over time, they realized that mortality is important, but you definitely have to also incorporate complications.

Co-Chair Bratzler: Any other comments from the committee? Okay, I believe Tristan, we can go ahead, and vote on the composite.

Mr. Wind: Voting is now open for Measure 3294 on composite quality construct, and rationale. The options are A for high, B for moderate, C for low, and D for insufficient. We are at 15 votes, last call for voting. Voting is now closed for Measure 3294 on composite quality construct, and rationale. There were six votes for high, nine votes for moderate, zero votes for low, and zero votes for insufficient. Therefore the measure passes on quality construct, and rationale.

Co-Chair Bratzler: All right, so we'll go on with scientific acceptability. Michael, if you could talk about reliability.

Dr. Firstenberg: Sure, I lost the -- so I'm sorry, the reliability is the next one?

Co-Chair Bratzler: Yeah.

Dr. Firstenberg: Sorry, I thought I was going on something else. But the overall reliability reflects that this is data that is collected more, or less in

real time submitted voluntarily to the STS where it's been validated, risk adjusted, and adapted in real time with comparisons of the different groups. Sorry, the different contributing programs ranging from 30 cases per year all the way up.

I think it's been consistent over time, it's been validated, the correlation coefficients show that it is, again, predictable, reliable, and reproducible.

Co-Chair Bratzler: Dr. Kertai?

Dr. Kertai: I agree with Michael. Measure specifications have not changed since the last review, and these specifications are very clear in design, so I don't have any concerns.

Dr. D'Agostino: I agree, the specifications are very well defined, and they absolutely can be consistently implemented, no concerns.

Co-Chair Bratzler: So, I don't know which member of our committee, somebody just highlighted the reliability scores that were in the submission, that they were in the 50 to 60 percent range. And of course I'm not exactly sure how that was measured. A little bit higher for larger groups, higher volume has higher reliability. I don't know, anyone from the measure developer wants to comment on that?

Dr. Jacobs: I think that the measure's reliability is certainly still within what's acceptable for performance of the measure, so no measure has perfect reliability, but certainly the numbers that we had, and the numbers that we documented in our submission are within what one would view as acceptable, and suitable for both feedback participants, and also for public reporting.

Co-Chair Bratzler: Anyone else on the committee have any questions, or comments? A quiet committee this afternoon. So, we'll move on to voting on reliability.

Mr. Wind: Thank you Dale. Voting is now open for Measure 3294 on reliability. The options are A for high, B for moderate, C for low, and D for insufficient. We are at 13 votes, last call for voting, last call. Voting is now closed for Measure 3294 on reliability. There were two votes for high, 11 votes for moderate, zero votes for low, and zero votes for insufficient. Therefore the measure passes on reliability.

Co-Chair Bratzler: All right, very good. So, validity, discussion of does the measure actually reflect quality of care given. Michael, validity?

Dr. Firstenberg: I think in brief, again, emphasizing we're looking at probably some of the most important quality metrics out there, and that is morbidity, and pretty well established -- excuse me, mortality, and pretty well established major morbidities that are associated with this procedure, and it's been consistent over time. And I think it is a very important, and robust series of metrics that programs can, and should be evaluated against.

And that the public can use to try to help determine where they want to get their care from, so I think it's extremely valid.

Co-Chair Bratzler: Dr. Kertai?

Dr. Kertai: I agree with Michael, and in addition to that, from an analytical sort of angle, in all six domains that I've listed under validity, the composite construct performed well, and any questions, concerns have been adequately addressed.

Dr. D'Agostino: I agree as well with Dr. Kertai.

Co-Chair Bratzler: Okay, I'm looking through the pre-submission comments, I don't see any great concerns here. I think it was pointed out in the comments that the composite score did correlate with actual patient outcomes such as mortality,

which seems to support validity of the metric. Anyone else, any other committee members?

Dr. Joseph: Yeah, this is Vilma. I just had two quick questions. One, I just wanted them to go over how they handled the missing values. They said they were imputed, so in terms of just getting the average, do you think that's sufficient for those certain variables? And the other one is that there was no social risk stratification that they documented.

I understand they said that they don't collect data on race, and ethnicity, but they do collect insurance, because we saw that in terms of the number of cases, but I was wondering what are they planning on doing in the future with regards to the risk stratification.

Dr. Shahian: I could address the STS issue if I could.

Co-Chair Bratzler: Sure, go ahead.

Dr. Shahian: We are in the process right now, in the STS database, of acquiring geocoded area depravation indices for all our patients. At least we've started first with our cardiac surgery guys, but we are hoping to expand this throughout the database. We believe that that area based indicator, and many of you on this conference are familiar with that, it comes out of the University of Wisconsin based on 17 STS variables.

We think that's the most comprehensive single STS indicator there is, and I hope that when we come back to you next cycle for re-endorsement, that we will actually have that ADI indicator incorporated.

Co-Chair Bratzler: Thanks David. What about, I want to make sure that we covered the question on just missing data. So, it's voluntarily collected, and submitted by the hospital systems that participate in the database, or the search.

Dr. Jacobs: I can make a few comments on that. First of all, all STS databases take missing data very seriously, and realize that the missing data can have impacts on our risk models. So, the first key is that in order for any hospital, or database participant to have their data included in STS aggregate data, receive feedback reports, and have their data included in our public reporting initiative.

That hospital is required to meet very stringent levels of data completeness for the key important fields, including operative mortality, which really means 30 day status, and discharge status. And if a hospital does not comply with those minimum standards of data completion, they're not even included in the aggregate data, and they don't receive feedback reports, and they cannot publicly report.

The number of hospitals that don't meet those requirements has decreased substantially year after year, after year. And now there's very few hospitals that do not have adequate completion of data, and very low rates of missing this on the key important fields. Beyond that, to more specifically address this question, during model development STS uses mechanisms of multiple imputation to deal with missing data.

I'm not a statistician, I'm a heart surgeon, but I do know that the statistician who works with us, who is a PhD statistician at Duke University named Shawn O'Brien is really a world leader on this type of methodology to deal with missing data, and he tells us that the multiple imputation methodologies used in all STS risk models are really state of the art, and the best way to handle missing data during model development. I think that's about as deep as I can go with this discussion.

Co-Chair Bratzler: Okay, thank you. Any other committee members have questions, or comments? All right, Tristan, I think we can move to voting on

validity.

Mr. Wind: Thank you Dale. Voting is now open for Measure 3294 on validity. The options are A for high, B for moderate, C for low, and D for insufficient. We are at 14 votes, last call for voting. Voting is now closed for Measure 3294 on validity. There were three votes for high, 11 votes for moderate, zero votes for low, and zero votes for insufficient, therefore the measure passes on validity.

Co-Chair Bratzler: All right Michael, next section is on the empiric analysis of the composite.

Dr. Firstenberg: Sure. So, with regards to that, the analysis is based upon, as we mentioned, the operative mortality, and major complication rate for the 186 hospitals that are contributing at least 30 lobectomies per program over the three year period, which is used to generate the risk adjusted model that then goes into the star rating system both for mortality, as well as major complications.

Has been discussed, this is all a component of the reporting system within the STS. There is appropriate weighting that is given to all of this, and I think that there is a very high level validity for this composite measure, and the empirical analysis to support how we're doing this.

Co-Chair Bratzler: Dr. Kertai?

Dr. Kertai: I can echo that, and I don't have any concerns about the empirical analysis for validity.

Dr. D'Agostino: I agree with that, I don't either.

Co-Chair Bratzler: All right, and I'm just looking through the pre-meeting comments, and it looks like there were no concerns reported. Anyone else on the committee have questions, or concerns about the metric? Seeing shaking heads no, so make sure there's no hands raised. Okay, we'll go

ahead with the vote. Tristan?

Mr. Wind: Thank you Dale. Voting is now open for Measure 3294 on empirical analysis. The options are A for high, B for moderate, C for low, and D for insufficient. There are 13 votes, last call for voting. Voting is now closed for Measure 3294 on empirical analysis. There were two votes for high, 12 votes for moderate, zero votes for low, and zero votes for insufficient. Therefore the measure passes on empirical analysis.

Co-Chair Bratzler: All right, thank you. Feasibility, Dr. Michael, do you want to talk about feasibility? This is voluntary data collection as I understand.

Dr. Firstenberg: I mean it is voluntary data collection, but I think it's for the most part something that we all embrace as being a core component of our professional obligation to the patients, and the public by reporting this. And I think people recognize that, as Dr. Jacobs alluded to, if you're not participating in it, then that does sort of reflect a little bit on your program.

And it is sort of a badge of honor so to speak, to be actively engaged. So, it is something that I think everybody takes pride in participating in, and trying to get the best numbers as possible, and part of a quality improvement program that we all need, if nothing else, for our maintenance, and certification as well. So, I think it's very strongly supported.

Co-Chair Bratzler: Dr. Kertai?

Dr. Kertai: I agree with Michael, it is very strongly supported within the thoracic surgery community, and thoracic surgeons take pride in participating in the process, and really helping to improve, and understand the ways they can improve the outcomes of their patients. So, I don't have any concerns about the feasibility.

Dr. D'Agostino: Neither do I, and I think it's also

important to point out that the data elements are routinely generated, and available electronically, and are well defined. So, I have no concerns about feasibility.

Co-Chair Bratzler: So, I just want to make sure, because it's been awhile since I've looked at the STS data set, and the variety of metrics that STS has. So, much of the data is available in electronic data fields, but I assume that still is transferred somehow into a separate reporting tool of some type.

Dr. Jacobs: So, the data is entered on site at a given hospital by data managers. The data managers have extensive training, including an annual data managers training course that's available both in person, and online, as well as monthly webinars that teach them about data quality, and data entry requirements. Those data are entered at the local hospital in the overwhelming majority of cases.

Then they're unloaded to our data warehouse, which is a company called IQVIA, and the data from the data warehouse then undergoes data analysis at the STS analytic centers. And once it's analyzed, it's then utilized to populate feedback reports, which are sent back to the hospital participants benchmarking their individual results against national aggregated data.

And it's also fed into the public reporting platform for those hospitals that choose to report. I'm not sure if that answered the question, I think it did.

Co-Chair Bratzler: Yeah, I think it did. Dr. Curry, you had a comment, question?

Member Curry: So, the challenge that we have in our health system is identifying those who have died out of hospital. So, how do these groups identify if someone has died within 30 days of the procedure, but out of hospital, so it would be a

manual process ---

(Simultaneous speaking.)

Member Curry: -- accuracy.

Dr. Jacobs: That's a great question, and it's been a big area of focus for the Society of Thoracic Surgeons across all of our databases. You're absolutely right that it's easier for a hospital to enter into a database whether, or not a patient was alive at the time of hospital discharge than it is to verify life status at 30 days if a patient was discharged home prior to 30 days.

You're absolutely right about that, and it's something that STS has focused on over the years, to develop methods to capture this data, and to ensure that it's complete, and accurate. We know it's accurate from our audit process, we have one of the largest audit processes of any registry on the planet, where ten percent of the sites are randomly audited on an annual basis.

And the audit includes validation, or verification of a number of important fields that are in our risk models, including obviously discharge status, and 30 day status. So, we know the data is accurate in those fields. In order to ensure it's complete, as I described before, hospitals are required to have 98 percent completion of the 30 day status field in order to have their data included in the aggregate STS data.

And to receive risk adjusted outcomes for their hospital, and that 98 percent was increased over a number of years from 90 percent, to 95 percent, to 98 percent. So, we've encouraged hospitals to be complete, because it's required for participation, and we encourage accuracy through our audit process. Beyond that, we require documentation of how 30 day life status was verified if a patient has been discharged home before the 30 days end.

So, one way is that there's proof of life in the electronic health record. Another way is that the patient comes back to see the surgeon for a follow up visit, that probably covers over 90 percent of our verification of life status at 30 days. But for the remaining ten percent, additional efforts are needed, which could include telephoning a referring doctor to check on the status of the patient.

Or even calling the patient themselves. For most cases those steps are unnecessary, because the patient comes back to see the surgeon for follow up, or there's proof of life in the electronic health record.

Dr. Shahian: I would just add to what Dr. Jacobs said, that we are looking for other methods as well to verify 30 day mortality. In the past it was possible to use the social security death master file, that's no longer feasible. But we have established a relationship which should allow us to use the national death index.

So again, by the time we come back to you for our next measure re-endorsement, we hopefully will be able to use national death index data. And in terms of manual, versus some other form of data entry, we believe right now that manual abstraction is the most accurate. But we also realize the data collection burden that imposes on programs.

So, actually STS has established a contractual relationship with a major university center to explore automated methods for extracting some of that element. Even if we could reduce data collection burden 30, or 40 percent, it would be major, as long as we don't sacrifice accuracy. So, again, that is an ongoing project that actually looks very promising.

Co-Chair Bratzler: Thanks Dr. Shahian. Any other comments, or questions from the committee? Seeing no more hands raised, all right Tristan, can we go ahead, and vote on feasibility?

Mr. Wind: Yes. Voting is now open for Measure 3294 on feasibility. The options are A for high, B for moderate, C for low, and D for insufficient. We are at 14 votes, last call for voting. Voting is now closed for Measure 3294 on feasibility. There were four votes for high, ten votes for moderate, zero votes for low, and zero votes for insufficient. Therefore the measure passes on feasibility.

Co-Chair Bratzler: All right, thank you. Michael, if you want to talk a little bit about use, how is the measure being used at this point? Did I lose you Michael?

Dr. Firstenberg: Sorry about that. This is used, and incorporated into just about every major thoracic program that's out there. Obviously there's some smaller programs that don't meet the volume criteria that may not necessarily get a reporting score, but that doesn't mean that they're not submitting. In fact probably every major program is using this extensively.

Both internally for their own quality improvement initiatives, as a component of all the STS data that they get, as well as those that are doing well, really emphasize the public reporting nature of it in terms of their program growth, and development. So, this is used extensively.

Co-Chair Bratzler: Dr. Kertai?

Dr. Kertai: I agree with Michael, this is used extensively in an accountable, and transparent way.

Co-Chair Bratzler: Dr. D'Agostino?

Ms. White: Dr. D'Agostino, I believe you're on mute, we're unable to hear you.

Dr. D'Agostino: Thank you. I was saying to myself that the GTSD, the thoracic database provides two reports a year, feedback reports to all of the participating sites, and that gives an immense

amount of information for quality improvement, and the results are publicly reported, so the public can easily look that up.

Co-Chair Bratzler: So, actually that was one of my questions. I noticed in the review it was listed as publicly reported. I assume that's predominantly voluntary at this point, in other words a center could participate, but not necessarily, or do you require it to be released publicly at this point?

Dr. Jacobs: The STS public reporting initiatives are all voluntary. I think the only entity that can require public reporting would be the government, and as a professional medical society, we can facilitate public reporting, and make it voluntary. That being said, we have very high rates of public reporting with close to 90 percent in our congenital heart surgery database, close to 85 percent in our adult heart surgery database.

And an increasing number every year of participants in the general thoracic surgery database. So, we continue to make efforts to encourage as many participants as possible to voluntarily publicly report, but in the absence of us being the government, instead with us being a professional medical society, that's what we can do, we can encourage our participants to publicly report.

Co-Chair Bratzler: Yeah, that was my kind of -- thank you. Others have questions on use, comments? I'm trying to scroll through, and make sure I've not missed raised hands here. Anyone else on the committee have any questions, or comments about use? All right, that being done, Tristan?

Mr. Wind: Voting is now open for Measure 3294 on use. The options are A for pass, and B for no pass. 13 votes, last call for voting. Voting is now closed for Measure 3294 on use. There were 14 votes for pass, and zero votes for no pass. Therefore the measure passes on use.

Co-Chair Bratzler: All right Michael, could you talk a little bit about usability, how do the centers get these reports, how are they using them?

Dr. Firstenberg: As has been mentioned, there's periodic reporting of the data that's been validated, and provides potentially, it would be risk adjusted outcomes of all these things that most programs incorporate into their ongoing continuous quality improvement projects that work hand in hand, both between the hospital, and the surgeon, and the surgical teams, and other disciplines.

And as you can see by the data that's been reported, that there is this ongoing reduction in mortality, and major morbidity that is essentially attributed to the fact that this data is being tracked, and reported, and does drive quality initiatives. Particularly with the emphasis on kind of the star rating system so that people feel that if they do embrace these things, they can get rewarded for being better beyond just the fact that they're helping the patients.

So, they see the data, and they can track their own institutional improvements. I think it's used extensively everywhere.

Co-Chair Bratzler: Dr. Kertai?

Dr. Kertai: Agree, the data, and the feedback, and the reports are used extensively in local, as well as national level type of quality improvement initiative programs related to the outcomes of thoracic surgery, including lobectomy. So, I don't have any specific incidence about usability.

Co-Chair Bratzler: Dr. D'Agostino?

Dr. D'Agostino: I agree with Dr. Kertai, I have no additional concerns.

Ms. White: Okay, and I'm looking through the pre-meeting comments, and essentially there were no

concerns, no unintended consequences that were identified, no harms. And some good news, that overall surgical operative mortality has decreased from 1.2 to 1 percent since 2015, so good news there. Anyone on the committee have other questions, or concerns? James?

Dr. Shahian: Yes, could the experts describe how the star system rating system is used, and at what levels it's used? Because I don't know exactly how performance translates into stars.

Dr. Jacobs: Yeah, I can try to tackle that one. So, first of all the statistical translation of performance into stars is based on Bayesian modeling, and this model used is utilizing 95 percent credible intervals, where outliers are true outliers with 95 percent Bayesian certainty. So, that's how one actually identifies if a program is a one star, two star, or a three star.

Once a program is identified as a one star, two star, three star, this can be tracked year, after year, after year to examine trends in performance. Three star programs can be studied to find out what are they doing that's good, and other programs can learn from these three star programs, and one star programs can do internal analyses to identify opportunities for improvement.

To learn strategies about what they need to do to become a two star program, because they've been identified as negatively performing outliers. Beyond that, STS is now in the process of creating a site visit program where expert surgeons from the society of thoracic surgeons will be available for consultation to visit any program to evaluate the program, and offer opportunities, and suggestions for quality improvement.

And this is a service that STS will provide to facilitate opportunities for one star programs to reach out to STS, ask for a site visit, and ask for advice on what that program can do to improve. But

I guess the one sentence answer to your question is that outliers are identified through statistical methodology which is sound, and therefore that allows identification of programs that are performing worse than expected, as expected, or better than expected.

Dr. Shahian: And just to expand a little bit on that, the underlying assumption to start with, the beginning in analysis is that all programs are statistically indistinguishable. Then we look to see whether the entire credible interval of a particular program is wholly above, or wholly below the STS average. Wholly above the scores, wholly above three star, wholly below one star.

And it is possible the way this is done, that we would have no one star programs if everybody was performing at a very high level. Unlike methodologies that use percentiles for example, there's no obligatory one, or three stars, and actually our goal would be to have every program in the country that's participating in this program be functioning at such a high level that we could not distinguish them.

And they would all be operating at an extraordinary level. We're not anywhere near that, but that would be our goal.

Co-Chair Bratzler: Any other questions, or comments from the committee on usability? All right, and I don't think I've missed anybody. So, Tristan, I think we can go ahead with the vote.

Mr. Wind: Thank you Dale. Voting is now open for Measure 3294 on usability. The options are A for high, B for moderate, C for low, and D for insufficient. We are at 14 votes, last call for voting. Voting is now closed for Measure 3294 on usability. There were six votes for high, eight votes for moderate, zero votes for low, and zero votes for insufficient. Therefore the measure passes on usability.

Co-Chair Bratzler: All right, thank you guys very much. So, that brings us to the last category, which is overall suitability for endorsement. I'll just open it up, and say does anybody have any comments. Michael, do you have any comments, or anyone else have any comments, or questions for the measure developers?

Dr. Firstenberg: This is a cornerstone for the thoracic oncology reporting system. I think it really sets the lead for just about everything else that's done in medicine, and surgery as kind of concurrent with, as Dr. Jacobs, and colleagues alluded to congenital heart disease reporting, as well as adult cardiac disease, kind of the three pillars of kind of pushing accountability within medicine to the forefront, that I think a lot of other disciplines are also working towards.

And it's -- stuff like this is without a doubt the most important reporting that we need to do. This needs to be endorsed unconditionally in my opinion.

Co-Chair Bratzler: I agree with Michael, and it has a downstream ramification, in fact also cardiac, thoracic, anesthesiologists working with thoracic surgeons, they work together, and sort of piggy back on the reports from the Society of Thoracic Surgeons, general thoracic surgery database reports on morbidity, and mortality, how to improve outcomes.

So, it's really a measure that draws quality, safety, and outcomes in the thoracic surgery patient population.

Co-Chair Bratzler: Any other comments, or questions about overall suitability for endorsement? All right, hearing none, Tristan, let's take the final vote.

Mr. Wind: Final vote. Voting is now open for Measure 3294 on overall suitability for endorsement. The options are A for yes, B for no.

We are at 15 votes, last call for voting. Voting is now closed for Measure 3294 on overall suitability for endorsement. There were 15 votes for yes, and zero votes for no. Therefore the standing committee recommends to endorse the measure.

Co-Chair Bratzler: All right, thank you guys very much. Thank you to all of our STS colleagues, really appreciate you being here today, and helping us understand this importance performance measure, and enjoyed working with you today.

Dr. Shahian: Thank you very much.

Dr. Jacobs: Yeah, thank you very much, and thanks for the service that you provide, you guys are doing important work.

Dr. Yagci: Thank you all.

Dr. Firstenberg: Thank you everybody.

Co-Chair Bratzler: So, I think I turn it over to LeeAnn at this point.

Ms. White: Thank you Dale. Thank you to the entire standing committee, and the developers for the great discussions we had today, and I know it can be a long day, but we did a great job reviewing these measures, and we definitely appreciate all the time that went into these measures. I have on the agenda a ten minute break between the end of our measure review period, and the beginning of our related, and competing discussion slides.

I will open it up to the standing committee. If you would like to take that ten minutes to stretch, and grab something to drink, or use the restroom, we can do that. Or we can move forward with the related, and competing, and the next steps. So, I will -- is there any objection to continuing on with the meeting?

Co-Chair Bratzler: Not from me, but I would prefer to get some time back at the end of the day.

Ms. White: Okay, so I will -- while Victoria is pulling up the slides, if I hear anything different, we'll just kind of power through, and go through the relating, and competing. So, I will wait for that moment to pull those up. Okay, all right, so I'm going to provide a brief overview of what is considered competing, and what is considered related. So, next slide Victoria please.

A competing measure is the same concept, and the same target population. In these instances, the standing committee would need to have a best in class discussion. We do not have any competing measures for this cycle. We also have related measures, where there are different target populations, or a different concept.

If they both are different, we don't have competition between those measures, and no harmonization is needed. If there are some similarities, developers are asked to harmonize their measures with the other related measure appropriately. Measure harmonization refers to the standardization of specifications for related measures with the same measure focus, or the same target population.

Or if the definition is applicable to many measures, so that they're uniform, and comparable, unless those differences are justified. The dimensions of harmonization can include numerators, denominator, exclusions, calculations, data force, and collection instruction. The extent of harmonization depends on the relationship of the measure.

Evidence for the specific measure focus, and differences in those data sources. So, next slide please. Before we begin the related review, I just want to note that this will not change the endorsement vote in any way, but the recommendations will be noted in the final report for future evaluation by the standing committee. So, we can discuss harmonization, and make those

recommendations. Next slide please.

So, our first related measure that we'll review is 3668, and the related Measure 3559, which is pediatric asthma emergency department use. Next slide please. We did put the details of 3599 up on the slide. The measure developer notes in their submission that 3668, follow up after emergency department visits for asthma is harmonized with Measure 3559 on age range, and with the ICD code used to identify the eligible population, and exclusion.

Next slide please. So, for this related measure, we would like the standing committee to consider these three questions. Are these measure specifications harmonized to the extent possible? Are there different effected impacts to interpretability, and add data collection burden? And are the differences justified?

So, I will pause a moment to open it up to the standing committee if there's any recommendation.

Co-Chair Bratzler: So, to open it up, I'll just start with just a couple of comments. I think although the target population is the same, the measure concepts are different. One is measuring population based use of the emergency department, the other measure is looking at whether the patient followed up after an emergency department visit for primary care, or specialty care visits.

So, I think overlap of the target population is quite appropriate, and I don't see them competing otherwise, my thoughts.

Ms. White: Okay, thank you. And anyone else from the standing committee, any feedback?

Member McCollister: This is Anna, I completely agree with Dale. I think that they're looking at the same population, same kind of care setting, but they're different measure issues.

Ms. White: Wonderful, thank you for that feedback.

Member Glomb: LeeAnn, this is Brendle, I agree with them, and it's a little bit skewed of a point, and purpose, but I don't see them directly competing with one another, and both probably useful.

Ms. White: Wonderful, thank you Brendle. Take a moment, it's hard for me to see the hand raises, I'm scrolling through the chat.

Co-Chair Bratzler: Yeah, me too, keep scrolling.

Ms. White: And I don't see anything -- Starlin agrees with Dale as well. So, I just wanted to pull that up for the record.

Co-Chair Bratzler: Any other comments from anyone about these two metrics?

Ms. White: And Kim agrees that they're different measures. Thank you so much, all right, we'll go to our next slide. So, this measure is 0061, comprehensive diabetes care blood pressure control 140 over 90, and this has been identified as a related measure to 0729, or optimal diabetes care measure. Next slide please.

So, in their submission, the developer notes that the two measure numerators are harmonized, and that there are differences in the denominator definitions due in part to the data source. The developer for 0061 uses claims data to identify diabetic patients, whereas Minnesota Community measurement uses EMR based data. The developer for 0061 uses methodology that looks for diabetic diagnosis codes, but additionally will include patients on oral medications, and insulin who do not have the diagnosis.

Lastly, Minnesota Community measurement notes that the diabetic women who are currently pregnant during the measurement year are excluded from measure 0729. So, I will pause here, again for

those questions that we have related to the measures, and see if anyone has any feedback, or recommendation for harmonization.

Member McCollister: If I remember correctly, we have a lot of competing diabetes measures, or at least measures related to the specific element of the composite measure. So, are we just taking this one measure at a time in terms of the measures that could be potentially competitive of the composite measure, or how is that working?

Ms. White: That's a great question Anna. So what we're doing is we're looking at the duty to measure, for example the 0729, 0061, they were identified as related, so they had the same target population, but different, and, or the same concept, or measure focus. And so we're looking at are these measures harmonized to the extent possible to reduce the burden of data collection, and use by accountable entities.

And if there are recommendations by the standing committee, we can note those in our final report. I hope that answers your question. Or is there additional questions on that? Right, these were not identified as competing, which would be the same concept, and the same target population.

Member Rosenzweig: These two measures, are they being curated by the NQF itself?

Co-Chair Bratzler: So, the other measure was submitted by the National Committee for Quality Assurance, and I believe that's already NQF endorsed.

Ms. White: Correct.

Member Rosenzweig: So, how did they get the data from -- all the complete data from claims? I don't understand how they get blood pressure, and so forth looking at the claims data.

Ms. White: So, for Measure 0061, we're not reviewing that measure, I don't have that measure specification in front of me, but that would be definitely something that would be in the developer's specifications for 0061. And that measure is currently NQF endorsed.

Member Rosenzweig: Okay, and Minnesota's, is unique in the sense that it has a very high use of electronic records, so that's why you could see that this is being used in Minnesota. The other thing is that both measures include patients on oral medications, and insulin, who do not have the diagnosis of diabetes. And also patients who are currently pregnant during the measure year.

Well, the problem with this -- this is the problem with the other measure, not the one that we're reviewing here, is that lots of -- type three diabetes is being treated with a lot of these medical agents, so you're really dealing with a population that includes a substantial amount of people that don't have diabetes, but just have pre-diabetes.

As I said, this is a problem with the other two measures, not for the measure we were analyzing. So, it's hard to harmonize, so to speak.

Dr. Jacobs: And I think this one specifically calls out blood pressure, which is a part of the composite measure, but we also have measures around A1C, and other elements that are part of the composite. So, I don't understand why this one is singled out as a competing measure, and those other measures are not.

Ms. White: So, this measure Anna, is identified as a related measure by the developer, and we did look at the NQF endorsement status for 0061 for the related measure. There was no other measures with the components that were identified as related. But I'm going to ask Matt Pickering to also weigh in on this. So, Matt if you can add more detail around the components, and the composite related.

Dr. Pickering: Yeah, so just for the composites, what you're looking at here is not only just a composite, if there's any sort of harmonization there that can happen at the composite level, but also the individual components. So, if you're looking at the components, if there's different patient populations, or different focus areas, is the measure harmonized to the extent possible?

And there could be justification as to why they're not harmonized to the extent possible -- sorry for the dog barking. It could be because it's a different care setting, it could be a different level of analysis, there could be some justification of why there's some differences. Whether you're looking at the blood pressure levels, or whatever they could be using some different data sources for that.

But the use may be different. They're looking at the composite score itself, but also looking at the individual components of those composites to see if there's harmonization with those measures, since they're related. And again, there may be a reason why they're not harmonized to the extent possible, because like I mentioned, the different care settings, or even different levels of analysis.

Member McCollister: Okay, I don't know if any of the others who have been on the committees for as long as I have recollection of some of the other specific components of this measure that have their own measures, but it feels kind of duplicative to me. I mean I haven't looked at the measures in a while, but I don't think the care settings are all that distinct. But anyway, we don't need to belabor the point.

Co-Chair Bratzler: Well, so Anna, I think I understand what you're saying. When you look at the components of the Minnesota measure, there are other performance measures that focus on some of those components, for instance percentage of patients with a hemoglobin A1C less than eight, or

less than either a variety of measures out there. Tobacco use, there are a number of tobacco use performance metrics out there.

So, the various pieces of this composite measure actually have other performance measures, that I believe some are likely NQF endorsed. I was trying to look for some of them, but I can't find them on the NQF website, but I'm sure they're there.

Member Rosenzweig: So, as I said, the difference is that these two other measures define diabetes differently.

And in the past there wasn't much of a difference, because we didn't use a lot of these agents like metformin, and some of the other agents for treatment of pre-diabetes, but now we do. And that means that a large percentage of the patients for these two measures, 61, and 59, a large percentage of them won't have diabetes as it's defined for our measure, the Measure 729.

So, it needs harmonization if you want to, otherwise it should not be considered harmonized.

Dr. Pickering: This is Matt Pickering from NQF. So, you had mentioned this a little bit earlier, you said that the concern really is about the other two, not this measure, 0729, but the other two measures may not be as expanded as they need to be to capture diabetes patients. Is that the concern?

Member Rosenzweig: Well, yes, so a lot of this measure, the denominator includes a lot of people who do not have diabetes. And the title of these two are under comprehensive diabetes care. So, we used to feel comfortable using the -- when this measure was created, it's a very old measure, 61, and 57. But when this measure was created, and this is part of our data set, when it was created back then, we used the use of diabetes medications as a way to define that the patient had diabetes.

And here, clearly, now that a large number of people are using certain medications to treat pre-diabetes, it means that there are different populations, and it's not really purely diabetes care. It's diabetes, and pre-diabetes care. I'd mention almost all people with pre-diabetes already have hemoglobin A1Cs under eight.

Co-Chair Bratzler: I guess what I want to point out is that NQF, you chose this measure because the measure developer listed it as potentially a competing measure, is that why we're talking about this one?

Ms. White: Correct, also in the measure submission, the developer identified 0061 as related, not competing, just as a related measure. So, again, these are the same population, and, or same concept. If it were the same target population, and concept, then we would look at it, and consider it as competing. So, this one was identified as a related measure.

And then they discuss how they harmonized the measures to the extent that's possible.

Co-Chair Bratzler: Okay, I guess the point I'm making, and I think the one that Anna's making is, so I am on the NQF website now, and NCQA has another measure, 0575, which is hemoglobin A1C control less than eight percent, I think that's part of the component that we reviewed today. Another one on LDLC control, and I'm probably not doing a really comprehensive search.

But there are other measures around other pieces of the component for this particular metric that I think are very similar, just like 0061 has some components of the measure that are similar to one piece of the composite measure we reviewed.

Member McCollister: Yeah, and I mean that's the point I was making (audio interference) the only -- the only actually competing measure when we've

got measures for all of the other components, is that -- I mean does NQF do a search of the measures, does the staff do a search of the measures to identify competing measures, do we just rely upon the developer to offer those up?

Dr. Pickering: It's both Anna, yeah. So, there's definitely what the developer has provided, and staff do try to present that to the standing committee, as well as other measures that could potentially be related, or competing. And then it's looking at those measures to determine if those measures come through as being passed, is there any further harmonization that needs to be considered?

And this would be recommendations both for the current measure that has come through, but also thinking about recommendations for the measures that are currently endorsed. Those other measures that are currently endorsed will come back through NQF maintenance endorsement, so capturing those recommendations within this forum would be helpful for the developer of those measures also to consider how they potentially harmonize further, especially with this measure that's coming through.

But there may be a need, or a reason why there's not fully similar numerators, and denominators. Like I mentioned that there is a different focus, or different care setting, or maybe even different intended use for the measure. There could be justification of why the numerator, denominator, or measure focus is a little bit different.

Co-Chair Bratzler: So, completely agree with that, I understand that. I'm just back to Anna's point, there are other measures that NQF has already endorsed that include other components of this composite. So, the developer gave you 0061, I would argue 0575, I think it's 0575 is similar to one component of the measure. I know NQF has a bunch of tobacco use, or tobacco cessation

performance metrics.

And I haven't even looked up Aspirin, I'm skimming past these, but there may be other measures, and again, as you pointed out there are very good reasons the denominator may be different, I get that, that's completely acceptable. Or the way the data is collected, how it's collected may be different, I get that. But there are other measures that overlap.

I think with components of this composite beyond just this one on blood pressure control.

Member McCollister: Yeah, it doesn't make logical sense to me, that we're just looking at this one measure with a competing measure -- Dale just listed off the numbers, but I know from sitting on the committee for more than a decade that we've got other measures that address that component. So, again, I don't mean to belabor the point, but it just doesn't make a lot logical sense.

I mean, more broadly speaking, I think it's important for NQF to ask to take a look at competing measures, and ensure that we're keeping them consistent, that we're not putting out confusing recommendations for quality measures. But I would also encourage NQF to think more broadly, and again, I'm not going to keep belaboring the point about that.

But if we're going to have measures where when we endorse these measures, we know that there will be consequences, for -- the adverse statin is 30 percent of the population by most recent estimates have some sort of adverse event. We need to come up with complimentary measures, but also require mitigation for the harms that are caused by the measures that we endorse.

Dr. Pickering: Yeah, and thank you very much Anna, and we definitely take that into consideration. The NQF obviously doesn't develop the measures,

so if there is a recommendation from the standing committee that there should be a balancing measure, or another measure that looks at trying to mitigate any unintended consequences, that is a recommendation that we can definitely provide within the report.

So, that future measure developers take that into consideration. And then if there's also recommendations from this group to look at the other components that you have stated, and have the developer ensure that there is harmonization to the extent possible for those other components, whether it be tobacco use, other things like that, that is the point of this related, and competing discussion.

So, that we can document those recommendations from this group, so that the developer can go back to providing further justification, or assessment on whether, or not they need to harmonize their measure even further.

Dr. Joseph: I have one comment to make. I was thinking about what you guys are talking about, looking for consequences, I don't want to incur a lot of burden for the developer, but maybe we could define certain clear clinical outcomes where they would know, okay you have requests for statins, give me your rate of allergic reactions, things like that.

And then you can track it, because it's easy for them to say no, there's no untoward consequences, but if we ask them to include it in their data collection. That may gave us an idea of are we really causing more harm than good?

Member McCollister: Yeah, it needs to be accounted for, thank you for that, it needs to be accounted for in some ways, and again, the latest data is that 30 percent of the population has an adverse event, and so if we're saying that to be a good quality provider, you have to meet this quality measure, or offer this

up as a measure that would suggest that you're a good quality provider.

It seems to me to be negligent to not require some sort of complimentary measure that ensures that people who choose this measure are actually doing the work to make sure that their patients are not part of the 30 percent of the population that has an adverse event. Again, our knowledge that this stuff has evolved.

And we need to take responsibility for ensuring that all of the elements acknowledged, and the evolution of our understanding of the impacts of these things are actually being included to make attempts to define, and enforce quality measures.

Dr. Pickering: So, I think that's great recommendations, and I know that we're sort of getting a little bit off the relating, and competing discussion, I wanted to kind of come back to that. But I think that this will be included in our meeting summary. I think there's some clear concern around wanting to see some unintended consequences related to some of the components of the measure, I believe statin use was mentioned.

So, we'll capture that information, it'll be something that will be included in our meeting summary, included in the technical report. This is something I think the developer would be made aware of, but is there anything related to this discussion on related, and competing that you would like to see the developer take a further look at?

Co-Chair Bratzler: Again, I have no problem with a composite measure. I think it's very useful in terms of thinking about comprehensive care. I will just in the chat here put down a few of the measures I'm finding that may have some overlap with components of the composite. They're separate measures, different developer, don't have any problem with separate measures, but there are measures other than just this one that they raised

in their submission.

Member Rosenzweig: For the two that they did raise however, I do think that the developers of the measure we've been discussing should not have to try to change their measure to try to harmonize it better with these two other measures. It would be better if the developers of those two measures harmonized better with the measure we're discussing.

Member McCollister: Why is that? I don't understand.

Member Rosenzweig: Because when they do their search for people that they define as having diabetes, they use a medication claims database, and anyone who's on a medicine that's used to treat diabetes is included. But in the last ten years there's been an increasing use of some of these medications to treat pre-diabetes, and pre-diabetes is very -- there are many more people with pre-diabetes than there are with diabetes.

So, that's a concern, so that a lot of the people who are being put into the denominator on these two measures, 61, and 59, really don't have diabetes, they have pre-diabetes.

Dr. Pickering: So, thanks for that --

Member McCollister: But there are people who are taking metformin for anti-cancer, anti-aging, I mean it's used a lot.

Member Rosenzweig: That's another issue that makes the use of claims data -- these medication claims data to define diabetes, it's just another issue as well. The EMR data doesn't do that, they basically look at the diagnosis.

Dr. Pickering: So, thanks for those comments. So, I think I'm hearing that the current measure, as it stands, James, based on your comments, that

maybe the other measures would need to have a look to see if there's some harmonization that needed to be done to maybe meet this measure, which in that instance, we'd have to wait for those measures to come back through for endorsement.

But part of what we ask developers to do is to take a look at other measures that currently are endorsed, and see how they could harmonize to those populations, that may be something that they're taking into account when they bring this measure back through. Which will be interesting to see if that is the case.

Related to this measure specifically, and Dale to your comments, is it fair to say that the committee is recommending that the developer take a look at the components in other NQF measures to see if there's harmonization efforts that have been made, and then if that's true, then we can definitely document that.

If there's no other comments from there, we can continue to move forward. Thanks Dale for dropping those measures into the chat.

Co-Chair Bratzler: There may be others, that's just one that I quickly found that's all diabetes related.

Dr. Pickering: Thank you. And I don't want to overlook the unintended consequences comments, I think those come through when we're actually evaluating this measure. It is something that we will capture within our meeting summary, and technical report for the developer's consideration, and we would also see what comes through with public comment related to any of that as well.

So, I appreciate those comments. Any other cements, or recommendations related to this measure, and other measures that -- yes.

Member Haydon-Greatting: Matt, this is Starlin, do you think they would be also dredging for the CPT

ICD9 for the statin induced myositis? Because it has its own -- I mean if I was pulling this data, I would include that as a subset pool of everybody that was pooled in the initial. Does that make sense? That helps answer Anna's concern.

That if they looked for people, within their pool of people that went to the doctor with muscle pain, and got the liver enzymes pooled, and that might be something that we could send as a recommendation to address that unintended consequence.

Member McCollister: I think that's a great suggestion, although I would say if the physician has attributed to statins, and identified that specific code, then they probably would then get to the -- they would be taken out of the population, because they would be identified as having an issue.

Probably what would be most helpful in identifying ones who have issues that have not yet been identified as being linked to statins is to look at the list of adverse events that happened in the 30 percent of the population that has adverse events related to statin. And then look for the ICD10 codes related to those symptoms.

As opposed to the one code that specifically says this person has statin related --

Member Haydon-Greatting: It's a series of events, what you'll see in the data is you'll see the start off slow, and then they move them up. And when they get to a certain level, if they need a high intensity statin, then they start exhibiting the side effects, then they back off. So, you'll see people back off, and do every other day just to mediate those side effects to make sure they can stay on the statin for good health.

But then you'll slowly see it just disappear off of their profile. But anyway, that's just something that we could maybe ask in their notes to look for that kind of stuff. I mean we've come a long way Anna,

from when we first looked at all this when we were all on the endocrine committee. Our data systems are so much more -- actually almost too big. But so there is ways to pull that kind of data.

Dr. Pickering: And just really on the unintended consequences piece, I want to note that this unintended consequences piece we can definitely capture, but that was something that we were reserving for the unintended consequences when we evaluate it for usability, which this committee had passed. So, we will make sure to capture this information, and some of the recommendations you had mentioned about looking in the data for certain CPT codes Starlin, that you had mentioned.

And some of the considerations you had mentioned as well Anna. I think that is something that we want to consider within the unintended consequences section of the usability component, just because that's evaluating if there is any unintended consequences with the use of the measure. So, we'll capture that as a recommendation for the developer to consider. But again, just kind of circling back here, on related, and competing discussions, and specifically related measures.

Outside of what Dale had recommended in looking at different components, different measures, was there anything else that the committee wanted to recommend for related, and competing discussions?

Related and Competing Measures

Member Rosenzweig: The problem with the CPT codes, to document things, is that they're often not used. I mean I see lots, and lots of patients who have some muscle aches, and muscle cramps, sometimes an elevated TPK, and I have to switch them to another statin. But I don't actually document in the CPT code that they have statin induced myositis. Because they've only been on it for a short time.

You switch them to another statin hat has fewer side effects, or you switch them to another drug that basically also lowers cholesterol.

Dr. Pickering: Great comment, thank you. Maybe something the developer has also considered when looking at unintended consequences, but as I stated, we'll make sure to document this in a report, but I think there were no other comments with related, and competing, so I think LeeAnn, we'll go to the next measure, and back to you.

Ms. White: Thank you Matt. Okay, so we are going to move onto -- next slide please Victoria. So, we have -- next slide. So, for 2797, we that's the transcranial Doppler ultrasonography screening among children with sickle cell anemia. There were two related measures that were identified, 3166 antibiotic prophylaxis among children with sickle cell anemia, and 3595, hydroxyurea use among children with sickle cell anemia.

Next slide please. So, here we have the measure description for 3166. So, this shows the description of the measure, percentage of children ages three months, to five years old with sickle cell anemia who were dispensed appropriate antibiotics, prophylactics, for at least 300 days within the measurement year. And then the next slide please for 3595, this is the percentage of children ages 1 to 18 years with sickle cell anemia who were dispensed hydroxyurea for at least 300 days within the measurement year.

So, the measure developer said both measure denominators include children with sickle cell anemia within the measurement year, measure 2797 is the number of children ages 2 through 15 years of age, whereas measure 3166 is the number of children ages three months to five years. For 3595, the children's ages range from 1 to 18 years of age with sickle cell anemia within the measurement year.

There's a cooperating note that all measures have been harmonized to the extent possible. Next slide please. So, the three questions again are the measure specifications, do the related measures harmonize to the extent possible, and are there any differences that could impact the interpretability, and add data collection burden? I will pause, and open it up to the standing committee.

Co-Chair Bratzler: Yeah, so this is Dale, I looked at all three before, there's some overlap in the denominator depending on measure, but otherwise they're all three measuring different aspects of preventive care for kids with sickle cell anemia, and just knowing the huge cost of that particular disease, particularly in the Medicaid population, I really can't figure out why more people aren't holding more providers accountable for some of these measures.

Ms. White: Thank you for that Dale. Anyone else from the standing committee have any feedback, or recommendations for harmonization among those measures? Okay, I'm just going to check real quick for any raised hands in the participant list. Okay, next slide please. So, I'll just take a brief moment, we are at the point of our measure eval meeting where we are opening it up for NQF member, and public comment.

NQF Member and Public Comment

So, I will provide a minute, or two to see if anyone from the participant list would like to come off mute, and provide their member, and public comment at this time. Okay, hearing none, I will now hand it over to our manager Isaac Sakyi, who will -- I'm sorry. We'll go through next steps, so next slide please. I will hand it over to Tristan, who will go through our overview of the remaining activities, and upcoming time lines for the project. So, Tristan?

Next Steps

Mr. Wind: Thank you LeeAnn. So, following today's meeting, staff will prepare the draft report detailing the standing committee's discussion, and recommendations. This report will go, and be released for a 30 day public, and member comment period. Staff will then collect the comments, and incorporate these into a comment brief, which will be shared with the standing committee, as well as the developers.

Following this, we will convene for our post comment call when the standing committee will reconvene to discuss the comments received. Following this call comments will be incorporated into the draft report to prepare for the CSAC meeting. And CSAC, the Consensus Standards Approval Committee, this is where they will meet to endorse measures, and following this, there will be an opportunity for appeals.

Next slide please. So, due to meeting all of our objectives today, we will be canceling the measure eval follow up meeting, so we will not be having that on June 28th. The draft report comment period will be held from August 3rd, to August 31st. And then CSAC review, and appeals period is to be determined, so we will communicate those dates as those come out, next slide please.

Here is our project contact info, if you are to have any questions, further comments, or concerns, you can email us, or call us at the following number along with the project page, and committee SharePoint sites. Next slide please. So, I'll turn it back to LeeAnn for outstanding questions, and conclusion.

Ms. White: Thank you Tristan. So, I will open it up for any questions at this time.

Member McCollister: I had a question, but I would say my mother happens to have issues with statin,

she's certainly not alone, but in terms of what I was saying about the need for sort of a companion measure to assess the adverse events associated with quality measures when we know going into the endorsement of the measure that there will be adverse events for some members of the population.

I think that's a broader issue that doesn't just apply to the statin measure, because we are learning, and have learned that a certain percentage of the population has adverse events to lots of different medications, and there are probably others related to more process measures that I don't know.

But I think it's something that needs to be considered, because we sit in these meetings, it's super easy, and I'm guilty of this as well, to overlook the real harm, and the frustration, and the life that adverse events -- or issues such as that cause in the life of individuals. And it's easy to think of them as a statistic, or just a number that may, or may not make its way to the denominator.

But there really needs to be careful consideration made for we're holding people to measures that are easy to check off, but not necessarily easy to follow up on, and make sure that there is no mark off. We need to get some serious consideration to steps that we can take to mitigate those concerns.

Ms. White: Absolutely, thank you Anna for sharing that, and I definitely appreciate all your feedback, and sharing that perspective. And I did write that down, we will note that on our end. And I think I agree, I feel it's a very important consideration. We should be looking at the unintended consequences, and I do value also the perspective that you all bring to the conversations.

And so, I will definitely take this back, we have noted it, and I will make sure to get this into the final report. So, thank you for providing that voice, and providing that feedback during these calls, it is

extremely valuable, and very important.

Member McCollister: Thank you.

Ms. White: Absolutely. Anyone else from the standing committee have any feedback, or questions regarding the measure evaluation meeting, or next steps? I'd be happy to answer.

Member Haydon-Greatting: I just want to thank all of you for all of your work, and support for all of us, so that we can get through this. And we have technical problems, and email problems, and life problems, and I also always want to thank Anna for bringing us to the reality of what our patients, and the persons with diabetes, or any sort of condition is going through, so thank you Anna.

And thank you NQF, everybody. And of course the co-chairs, what would we do without our wonderful co-chairs that keep us in the boat? So thank you.

Ms. White: Absolutely. Well, that's a great headway to my next slide, which is wrap up. But I want to echo Starlin's appreciation. So, I want to say thank you to the entire participant, attendees today. Big shout out to the standing committee, our wonderful co-chairs. I know Adam had to slip away at two for another commitment, but we definitely appreciate everyone's patience, and their engagement, and their participation.

I know not too long ago we had a fall '21 measure evaluation meeting, and post comment meeting, so I definitely wanted to extend that appreciation to you all. I know a lot goes into the measure review, and the pre-evaluation, and the meeting. So, thank you so much. I also want to thank our developers, if they're on the line, a big thank you to our developers for the time, and effort leading up to the meeting.

And attending the call today to address any questions, and concerns that the standing

committee had. And I also definitely want to thank my team, my entire team. So, Isaac, Tristan, Taroon, Matt, Victoria, Matilda, this is a wonderful group that I enjoy working with every day, and they work very hard. So, I want to definitely give them a thank you.

So, I hope everyone stays well, stays healthy, and stays safe. I hope the weather, and the sun start shining for you all. I know it's been a long day. So, we are ahead of schedule, and so we can give you back that hour, and a half today. So, if you need anything, please let me know, please email our project team, we're always here to assist you, and support you. I'm going to leave -- Dale, please provide your closing remarks, I'm sorry.

Co-Chair Bratzler: Nothing other than a great committee to work with, really enjoyed all the -- I always learn so much from the committee, so really appreciate everybody.

Adjourn

Ms. White: Okay everyone, thank you so much, let us know if you need anything, and until then please be safe, and stay well. Thank you.

Co-Chair Bratzler: Thank you.

(Whereupon, the above-entitled matter went off the record at 3:27 p.m.)