National Quality Forum
Patient Safety Measure Evaluation Web Meeting
Spring 2021 Cycle
Thursday, June 24, 2021

The Committee met via Videoconference, at 10:00 a.m. EDT, Ed Septimus and Iona Thraen, Co-Chairs, presiding.

Present:

Ed Septimus, MD, Co-Chair
Iona Thraen, PhD, ACSW, Co-Chair
Joel Bundy, MD, FACP, FASN, CPE, Sentara Healthcare
Elissa Charbonneau, DO, MS, Encompass Health Corporation
Curtis Collins, PharmD, MS, St. Joseph Mercy Health System
Theresa Edelstein, MPH, LNHA, New Jersey Hospital Association
Terry Fairbanks, MD, MS, FACEP, MedStar Health
Jason Falvey, PT, DPT, PhD, University of Maryland School of Medicine
Sara Hawkins, PhD, RN, CPPS, Eastern Idaho Regional Medical Center
Bret Jackson, The Economic Alliance for Michigan
John James, PhD, Patient Safety America
Raquel Mayne, MS, MPH, RN, Hospital for Special Surgery
Anne Myrka, RPh, MAT, Island Peer Review Organization
Edward Pollak, MD, Henry Ford Health System
Jamie Roney, DNP, NPD-BC, CCRN-K, Covenant Health System
David Seidenwurm, MD, FACR, Sutter Health
Geeta Sood, MD, ScM, The Society for Healthcare Epidemiology of America
David Stockwell, MD, MBA, Johns Hopkins University
Donald Yealy, MD, FACEP, University of Pittsburgh Department of Emergency Medicine
Yanling Yu, PhD, Washington Advocate for Patient Safety

NQF Staff:

Tamara Funk, Manager
Michael Haynie, Senior Managing Director
Yemsrach Kidane, Project Manager
Chris Millet
Matthew Pickering, Senior Director
Isaac Sakyi, Senior Analyst

Also Present:

Rachel Karen Campos, American College of Radiology
Bo Feng, PhD, IMPAQ International Katy Grainger
Kendall Hall, MD, IMPAQ International
Tom Heymann, Sepsis Alliance
Cheng Lin, Acumen, LLC
Mahadevappa Mahesh, PhD, Johns Hopkins Bloomberg School of Public Health
Sri Nagavarapu, Acumen, LLC
Emanuel Rivers, MD, MPH, FACEP, Henry Ford Hospital
Patrick Romano, MD, MPH, FACP, FAAP, IMPAQ International
Steven Simpson, MD, University of Kansas
Rebecca Smith-Bindman, MD, University of California, San Francisco
Sean Townsend, MD, Institute for Healthcare Improvement, California Pacific Medical Center
Chana West, RN, MSN, IMPAQ International
Dana Wollins, DrPH, Infectious Diseases Society of America

Contents

Proceedings

(10:06 a.m.)

Welcome and Review of Meeting Objectives

Mr. Pickering: So I want to welcome everyone to the Patient Safety Spring 2021 Measure Evaluation Cycle and thank everyone for your time, both the developers as well as the Standing Committee Members. Leading up to today's proceedings, as you know there's a lot of time that goes into developing these measures and submitting them for NQF endorsement as well as the time to review those measures, both prior to the meeting today but also during the meeting as well, so I want to very much thank everyone for your time and efforts, both for today and leading up to today.

My name is Matthew Pickering. You all on the Standing Committee know my face fairly well. I'm Senior Director here at NQF working with this great team here and also this great Standing Committee.

We'll go to the next slide, Isaac. In addition, I want to give an opportunity as well to our Co-Chairs, Iona and Ed, to provide some welcoming remarks for the proceedings today.

Iona, I'll start with you and then we'll go to Ed.

Co-Chair Thraen: Well, this was a way to wake my brain up, I just have to say, this morning. Welcome everybody. We'll get this coordinated. I just want to thank everybody for their time and their commitment and to let everybody know that they were having troubles getting in and staying, too, so thank you very much.

(Simultaneous speaking.)

Mr. Pickering: And before we go to Ed --

Co-Chair Thraen: Feedback.

Mr. Pickering: There we go, we got it. Thanks.

Co-Chair Thraen: Okay. And then I just want ---

Mr. Pickering: Iona --- go ahead.

Co-Chair Thraen: I just want to point out that there are two future chairs -- co-chairs -- on this meeting, Don Yealy and John James, and just acknowledge their involvement from this last year in terms of attending our work meetings behind the scenes and they will be taking over for Ed and I, and I want to thank them for that. Now, I'm done.

Mr. Pickering: Thanks, Iona. Ed?

Co-Chair Septimus: Also welcome everyone. I know it was very frustrating getting in this morning. I'll second Iona's stuff and, of course, welcome our new incoming Chairs who are going to be outstanding.

We have six measures for today. Last time, we actually got through in one day. We do have an extra half day tomorrow if we don't finish today. I'll just remind the Committee that if we stay focused and disciplined, I think we can finish today. Try to kind of keep track of what we're talking about per measure. So, if we're talking about the evidence, don't start talking about validity and reliability. It's so, so easy to slip into that, but let's stay focused to the element of the measure we're speaking about and then reserve comments when the reliability or validity or usability comes up. So let's stay focused in that regard. Be respectful of one another. If somebody else mentions what you were going to say, then it's perfectly okay to put your hand down to make sure we use our time effectively.

I just want to again thank everybody because this is a great committee, we're going to have a very

robust discussion, I'm sure, over the next day. And thank again NQF staff, who have worked tirelessly to kind of keep us up to date. There have been some changes in the staff, which many of you are aware of, but despite that, the NQF staff has stepped up and helped Iona and I and the whole committee make this day possible. So with that, I'll turn it back to Matt.

Mr. Pickering: Thank you, Ed, for the kind words. Again, I want to echo the thanks for everyone as well and the acknowledgment of John James and Don Yealy, who will be stepping in as co-chairs going into 2022, so thank you for your involvement in those proceedings that we've done with our co-chairs.

All right, so we do have -- as Ed mentioned, this meeting today is really spread out across two days so we do have some time tomorrow in case we're not able to finish all the proceedings today. So, that time will be reserved for tomorrow, that is on your calendar and there is another agenda as well, the day two agenda, for that meeting tomorrow. But for today, we'll see how far we can go.

Within the day one agenda, you see there are some asterisk measures in there that we are anticipating -- given the time constraints or how much discussion happens with the previous measures -- those would be moved to tomorrow's proceedings as 3501e and 3389. Those are the two measures we've anticipated would be moved to tomorrow, but the developers are going to be on the call today. We do appreciate your attendance and participation and patience as we get through all the material today, but those may be moved to tomorrow depending on how much time we have to go through today.

So, going to the next slide, I just want to touch on some housekeeping reminders. Again, this is a new

platform we're using for this cycle. It's the Webex platform, but it has all the same features as we've used in previous platforms. RingCentral was the most recent.

So, we have a raise hand feature. If you're not familiar with where that is, if you go to the participants list and you hover over your name, there is a little raised hand. Your name should be the very top name in the participants list, in case you're trying to find yourself. If you just hover over your name, there's a little hand that pops up there where you can definite -- and I see people already using that functionality, so it sounds like you found it. That's great. So, we will monitor that and just make sure that we try to recognize you as well.

We do have a chat feature, so please if you feel if you want to chat to everyone or want to chat to a certain member of the Committee or staff, you can do so within that feature. We'll try to monitor that as well and recognize you.

We always welcome you to participate verbally, but just adding to the discussion and dialogue where you can. If you are not speaking, however, we kindly ask yourself to keep yourself on mute, just to prevent any background noise. If you are speaking, we have a video feature. We love to see your face. It allows us to be a little bit more engaged here obviously. Maybe a little bit challenging if you're just using your phone, but we do encourage it. We'd love to see your faces talking during that part where you like to participate.

If you have any challenges, obviously you can use this chat feature and such to mention those challenges with any technology issues or you can e-mail the project box, which is patientsafety@qualityforum.org. Next slide, please.

So here is the team, as you see listed here, on the slide. We'll just do a brief intro here. So, Matt Pickering, Senior Director here at NQF, working on this project. Shalema Brooks. She's not on the call today. She is a new director. She is here to help support the project as well, but she is unfortunately not able to make the call today; however, she sends her welcome and thanks for all of your work leading up to this meeting.

I'll turn it to Tami to give some -- to say hi to everyone and then we'll kind of go down the list. So, Tami.

Introductions and Disclosures of Interest

Ms. Funk: Good morning. I'm happy to be here with you all today. Tami Funk and this is my second meeting with your committee at this point, so happy to join.

Mr. Pickering: Thanks, Tami. Isaac?

Mr. Sakyi: Good morning, this is Isaac Sakyi. I'm the senior analyst on this project and I'm looking forward to today's meeting.

Mr. Pickering: Thanks, Isaac. Yemi, if you're on?

Ms. Kidane: Good morning, everyone. This is Yemi Kidane. I am the projector manager on this project.

Mr. Pickering: Thanks, Yemi. And Jesse Pines may attend this meeting if he can. I know he's tied up at the clinic today. So, he may attend this meeting if he's able to do so, so Jesse Pines, who's a consultant, who a lot of you are very familiar with. This is the team, so we'll be running the back end work throughout the proceedings today. So I just wanted to recognize them and all the hard work that they've done. So, we'll go to the next slide.

Just for the agenda today, we'll first do introductions and disclosures of interest. That will be with Michael Haynie, who's our Senior Managing Director here within Quality Measurements at NQF.

That will be where there's an opportunity to provide some introductions for us to see who's present on the call. It allows us to also see if we have established quorum. Again, quorum is very important for us to establish in order for us to vote on the measures today. So, we appreciate you being in attendance.

During the disclosures of interest, there have been a few of you that have indicated that you have some potential conflicts with the measures that will be reviewed today. If you have indicated those, we kindly ask that you will state what that conflict -- that you have a potential conflict and what that is and that will potentially determine whether you're recused from the measure proceedings today. That recusal means that you will not be voting on the measure that you have a conflict with. Also, you will not be able to participate in the discussions as well. So just to note that a recusal, you would not vote -- so you would not vote on that measure, on any of the criteria for that measure -- and you will not participate in the discussions due to the potential conflict that you have identified.

After that, we will then go through an overview of the evaluation process and voting procedures and do a voting test. So I will pause here. There is always that challenge potentially of getting into the voting software. We did send out a Poll Everywhere link. It is a Poll Everywhere link and you've see this previously before. So, feel free to find that and go in there and start getting that ready to go for our voting test, which will just be a test question just to make sure we have everybody up and running. So, we will revisit that as we get to the voting test

portion of today.

Then, we will go through an introduction of -- or an overview of the measures under review today. Then, we will go into the actual evaluation and consideration of those measures for endorsement. I will talk about a little bit of the process for that as we get to that point.

If we are able to complete all of the measure evaluation proceedings today, we will do a related and competing measure discussion. If we are not able to complete that, we will reserve that for tomorrow. And then after related and competing, we will then have NQF member and public comment, but even if we didn't get to related and completing today, we still have space at the end of our meeting today to have NQF member and public comment. So we still will have time for those members of the public and NQF members as well, even if we didn't finish all of the measures today, we'll leave some space for anyone from the public to comment on the measures that have gone through endorsement consideration today.

Then we will end with next steps and just talk about the -- what happens next with all of the work that has happened today. If we have a meeting tomorrow, we will also do the same as well tomorrow.

Any questions about the agenda or what we'll be doing today before I turn it over to Michael Haynie?

Looking at no hands, no questions in the box. Okay. I will turn it over to Michael Haynie to do the introductions and disclosures of interest. Michael?

Ms. Haynie: Great, good morning everyone. Thank you so much for being here. So, we're going to combine DOIs and introductions and we're also

going to talk about recusals here since we have a smattering of those, so please bear with me, we've got a lot of material to go through to make sure we're all above board and transparent.

So, you have received from us at some point in time two disclosure of interest forms, right? So, one is our annual disclosure of interest --- you probably got that earlier in the calendar year -- and the other is disclosures specific to the measures we are reviewing in this cycle. In those forms we ask you a number of questions about your professional activities. I'd like to thank each of you. You all did a really good job on those forms, you got them back, they were very transparent. So because of that we do have some recusals and we will go through that as we do the introductions.

Today, we will ask you to verbally disclose any information you provided on either of those forms that you believe is relevant to this Committee. We're especially interested in grants, research or consulting related to this Committee's work.

As a couple of other fun and exciting reminders, you do 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 both paid and unpaid activities that are relevant to the work in front of you. And finally, in most cases, just because you disclose does not mean that you have a conflict of interest. We do verbal disclosures in the spirit of openness and transparency. So, what we're going to do here. We do have one, two three, four, five --- we have five actual recusals and disclosures, so we will cover those as we go through.

But what I'll ask you to do is we'll go around our virtual table here. I'll call your name, please state your name and correct me on pronunciation if I get

it wrong. I'm trying to improve on all of these. I'll ask you to state the organization that you're with and if you have anything to disclose and for the five of you who do have an actual recusal, I'm going to ask you to state that as well and if you don't know the measure number, the team has helpfully given me a list of the measure numbers so we can make sure we're above board.

As a general note, for those listening and sort of tracking, when we have someone who is recused, they come out of both the numerator and the denominator as we track both quorum and as we track those important numbers to get to for consensus. So, you will see some change in the numbers today and we will do our best to call that out and let everyone know so no one gets hung up or confused on the math. It is going to change a little bit measure by measure as we do work through these recusals, so that's not a surprise.

To kick us off, Ed, may I start with you?

Co-Chair Septimus: Sure, Michael, take a deep breath first, Michael. Deep breath.

(Laughter.)

Co-Chair Septimus: I am present obviously and I am going to be recused from the first measure, 0500. I'm on the IDSA Task Force on Sepsis and a co-author of the ISDA physicians paper on sepsis. I also did provide several years ago a few Mathematica, some consulting, on the surviving sepsis measure regarding anticoagulation and blood culture contamination, so I am recusing myself from Measure 0500.

Ms. Haynie: All right. Thank you for that. Iona?

Co-Chair Thraen: Iona Thraen from the University of Utah Health Systems. I have no conflicts of interest.

Ms. Haynie: Thank you.

Co-Chair Septimus: Oh, I forgot. I forgot to tell you where I'm from, Harvard Medical School. Sorry.

Ms. Haynie: Thank you, Ed. We'll skip Emily Aaronson since she's inactive at the moment. Joel Bundy?

Member Bundy: Good morning. Joel Bundy present from Sentara Healthcare and I have nothing to disclose.

Ms. Haynie: Thank you. Elissa Charbonneau?

Member Charbonneau: Good morning, everyone. I am present, Elissa Charbonneau, from Encompass Health. I have nothing to disclose.

Ms. Haynie: Thank you. Curtis Collins?

Member Collins: Yeah, good morning, everyone. Curtis Collins, St. Joseph Mercy Health System and I have nothing to disclose.

Ms. Haynie: Thank you. Theresa Edelstein?

Member Edelstein: Good morning, everyone. I am from the New Jersey Hospital Association and I have nothing to disclose.

Ms. Haynie: Thank you. Jason Falvey?

Member Falvey: Good morning, everybody. Jason Falvey. I'm faculty at the University of Maryland School of Medicine and I don't have anything to disclose.

Ms. Haynie: Thank you. Jason, just so you know, I can hear you, but your audio is pretty faint, so if you're trying to get a word in later, be aware that your sound is kind of low right now. Terry Fairbanks?

Member Fairbanks: Hi, everyone. Terry Fairbanks. I'm Vice President of Quality and Safety at MedStar Health and a professor of emergency medicine at Georgetown University and I have no conflicts to disclose.

Ms. Haynie: Thank you. Robert Green? Sara Hawkins?

Member Hawkins: Good morning, it's Sara Hawkins from Eastern Idaho Regional Medical Center, part of HCA, and I have no disclosures.

Ms. Haynie: Thank you. Bret Jackson?

Member Jackson: Bret Jackson with The Economic Alliance for Michigan and I have no conflicts to disclose.

Ms. Haynie: Thank you. John James?

Member James: John James, Patient Safety America, where I'm the founder and CEO, in Houston and I have nothing to disclose.

Ms. Haynie: Thank you. Laura Kinney? Arpana Mathur? Raquel Mayne?

Member Mayne: Raquel Mayne, Hospital for Special Surgery. I have nothing to disclose.

Ms. Haynie: Thank you. Anne Myrka?

Member Myrka: Hello, my organization is IPRO and I have nothing to disclose.

Ms. Haynie: Thank you. Edward Pollak?

Member Pollak: Good morning. Ed Pollak, Henry Ford Health System and I am recusing myself from 500, the sepsis measure, because Henry Ford Health System is the measure developer.

Ms. Haynie: Thank you for that.

Member Pollak: Thank you.

Ms. Haynie: Jamie Roney?

Member Roney: Hi, I'm Jamie Roney and I work at Covenant Health, part of Providence St. Joseph Health System, and I do not have any conflicts of interest to disclose.

Ms. Haynie: Thank you. Nancy Schoenborn? David Seidenwurm?

Member Seidenwurm: Yeah, hi. I'm David Seidenwurm. I'm a neuroradiologist in Sacramento, California. I work at Sutter Health and I'm the Medical Director for Quality and Safety in the network, Medical Director for our managed care plans. And I have a conflict with respect to I think it's Measure 3621 because I'm Chair of the Metrics Committee for the American College of Radiology and also the principal investigator on The Moore Foundation Diagnostic Excellence Grant through the ACR and a measure developer with the ACR.

Ms. Haynie: All right. Thank you for that disclosure. Geeta Sood?

Member Sood: Hello, I filled out that I did not have disclosures, but now I have a question. Obviously the 500 metric has a lot of strong opinions and on behalf of SHEA as the Chair of the Quality Metrics Task Force, I will be putting together comments and I have published a little bit in the area on unintended consequences, do I --- I hope not, but do I need to recuse myself?

Ms. Haynie: So, Geeta, generally -- that's a great thing to disclose. We appreciate that. Typically, when we see something that would be a recusal, it's that you have participated in the measure

development in some way, been on a panel, done some face validity testing, done that, so I just want to be clear of -- or as a measure developer that's another big one. So, if you've just published opinions or had opinions, that is a great thing to disclose, but does not generally trigger a recusal.

Member Sood: Excellent.

Ms. Haynie: So with that information, how are you feeling about that?

Member Sood: That sounds great and I don't think I would meet of any those criteria which is why initially I hadn't put anything down for the disclosure. So perfect, thank you.

Ms. Haynie: Okay, great. Appreciate the transparency. All right, David Stockwell?

Member Stockwell: Hello all. David Stockwell. I'm a pediatric intensivist, the Chief Medical Officer at Johns Hopkins Children's Center in Baltimore. I also work with Pascal Metrics, which leads to the disclosure that that organization, in the very early days of the 3501e, that they were essentially a data gathering entity for the measure developer. And so for that reason, I will be recusing myself from that measure.

Ms. Haynie: Okay. Thank you for that. Donald Yealy?

Member Yealy: Hi, I'm Don Yealy. I'm a distinguished professor and Chair at the University of Pittsburgh and recently became Chief Medical Officer for the UPMC health system. I am recusing from Measure 0500 in that I have completed one and have an ongoing active federally funded trial in evaluation and treatment of septic shock. I am the first author on American College of Emergency Physicians sponsored consensus comments on

current guidelines and I have had an active role in previous 0500 testimony and the SEP-1 measure from CMS that emanated from 0500, so I'll recuse.

Ms. Haynie: All right. Thank you for that. Yanling Yu?

Member Yu: Yeah, good morning. I'm with the Washington Advocate for Patient Safety and I have nothing to disclose.

Ms. Haynie: Thank you. All right, so while our staff is counting all our different variations of quorum today, I have some more information to share with you.

First of all, thank you all for that and for the spirit of transparency and openness. I would like to let you know that if you believe you might have a conflict of interest at any time during the meetings as topics are discussed -- sometimes that happens. We get into something and you're like, oh wait, I worked on that and didn't realize it was going to come up here. Kind of like Geeta just did. Geeta provided some great modeling for us, so you may do what Geeta did and ask a question in chat. You can also talk to the Co-Chairs. I would remind you one of our Co-Chairs is recused from 0500 so please talk to Iona about that one if you have questions. Or, you can talk to NQF staff. We are here and available for you. If you believe that a fellow committee member may have a conflict of interest or is behaving in a biased manner, you may also point that out during the meeting by sending a message to your Co-Chairs or NQF staff.

Does anyone have any questions or anything else you'd like to discuss based upon the disclosures made today?

Co-Chair Septimus: Mine's got nothing to do with

that, but so we have, by my count, we had four people that may not be here, is that correct?

Mr. Pickering: So, yes, Ed, you are spot on. Yes, we have four people.

Co-Chair Septimus: So there's four that's not, we have -- I'm just trying to make sure. I don't want to run into problems with counting later -- and Matt knows what I'm talking about. So, we have four that are not here and for the first measure, we have three that are recused, so that means the quorum has to be 14, so that means we still have, my count, 22 members that can vote on the first measure. Am I getting that correct?

Mr. Pickering: Ed, you know, I don't even need to be here. I can just go, Ed. I can leave. You can do it.

Ms. Haynie: You do, you do need to be here, Matt.

(Laughter.)

Co-Chair Septimus: I just wanted to make sure. Trying to keep track of the numbers. Matt knows why I am doing that.

Mr. Pickering: You are correct, Ed. And so I will just state -- and Michael, I think we were going to transition. With one inactive member that drops this down to 24 people that are active. So, those are active people that are eligible to vote.

Co-Chair Septimus: That's it.

Mr. Pickering: Not including the recusals, right? So, we have 20 people on the call, with the recusals now in place for those measures -- so for example 0500, right? We have three recusals, so a quorum for that would be 14 people. We have 20 people on the call, so we have met quorum. Right? We would

actually need 17 because of those three recusals, right? So, we actually are above that, we're at 20. So, we are good for quorum overall and we're good for quorum for all of the subsequent measures with recusals. That's good.

Co-Chair Septimus: Thank you, Matt. The other thing is that occasionally members of the Committee may need to leave for short periods of time. If you'll let us know that so when it comes to voting we know who's present and who's not.

Ms. Haynie: Great. All right and we will keep track of this if you need to step away, chat us, let us know. It remains to me to just sort of remind you all that NQF is a non-partisan organization, so out of mutual respect for each other, we do encourage that everyone make an effort to refrain from making any comments, innuendos or humor relating to, for example, race, gender, politics, other extremely polarizing topics that otherwise may be considered inappropriate during the meeting.

We encourage discussions that are open, constructive and collaborative. And especially, I know many people have talked about the sepsis measures, we expect to have some deep and introspective discussions on this, so let's be sure that we are all mindful of how our language and opinions may be perceived to others and we stick to the criteria and the topic here. NQF staff is here for you to get back on track and sort of make sure that we do refocus ourselves on the algorithms and the criteria today.

To Bret's point that I saw come up in the chat, sometimes while we're doing introductions people do join. Did I miss anyone who is part of the Standing Committee who did not get an opportunity to introduce themselves and disclose?

All right. So, I will turn it back over to the team.

Mr. Pickering: Great. I just want to thank everyone once again. Next, we will be going over the evaluation process. I will turn it --- just looking at the chat. I will turn it over to Tami, who will go through the overview evaluation process. So, thank you. Tami?

Ms. Funk: Thank you, Matt. I'm assuming everyone can hear me. Please shake your head if no sound is coming out. Okay.

Mr. Pickering: You sound good.

Overview of Evaluation Process and Voting Process

Ms. Funk: Thank you. So, the role of the Standing Committee during today's evaluation meeting is to act as a proxy for our NQF multi-stakeholder membership. The Committee will be evaluating each measure against each criterion and providing a rating to represent the extent to which each criterion is met and their rationale for that rating. The Committee will also make recommendations regarding endorsement and will be responding to any public comments and, in general, the Committee oversees the portfolio of patient safety measures. Next side, please.

On this slide, we have some ground rules for today's meeting. In general, the Committee should have already been prepared and reviewed the measures before this meeting. We ask that you please remain in the meeting at all times and if you have to leave early, please let us know. Please remain engaged, but also allowing others to contribute and keep your comments concise and focused. Does everyone agree to these ground rules today? I see heads nodding so we'll take no objections as agreement.

Our process for measure discussion and voting is

that the measure developer will introduce their measure and then the Committee members who are lead discussants will begin discussion of the measures in relation to the evaluation criteria. Each criterion will be discussed and then voted on before moving to the next criterion. During the introduction and discussion of measures, developers may respond to questions from the Committee and correct any misunderstandings about their measures. However, the discussion surrounding the evaluation of measures is primarily meant for the Committee. Measure developers will be invited to respond to questions at the discretion of the Committee.

During the evaluation discussions, Committee members may offer suggestions for improvement to measures and these can be considered by the developer for future improvements; however, the Committee is expected to evaluate and make recommendations on the measures using the submitted specifications and testing only. Remember that as Committee members, you act as a proxy for NQF's membership. As such, this multi-stakeholder group brings varied perspectives, values and priorities to these discussions. Please respect differences of opinion and we ask for collegial interactions among Committee members and measure developers. Next slide, please. Thank you.

So the next three slides detail NQF's endorsement criteria. There are five categories of criteria that we will be discussing today. Importance to measure and report, scientific acceptability, feasibility, usability and use, and comparison to related or competing measures. Next slide.

Votes will be taken after the discussion of each criterion. The Committee will first review and vote on evidence and performance gap, which are must

pass criteria, followed by reliability and validity, which are also must pass and then feasibility, usability and use, which is a must pass criterion for maintenance measures.

Please note that the voting platform will capture your most recent vote, but we have discovered some nuances to updates in our voting platform, so we may be asking you to clear your last vote and revote in the moment so that we make sure we have an accurate count and that we know that our denominator is accurate in the moment.

Again, please chat us if you're having any issues with that voting platform or have any questions. Next slide, please.

Each measure will also receive a discussion around related and competing measures. Provided that a measure passes all must pass criteria, it will be voted on for overall suitability for endorsement. If a measure fails any must pass criteria, discussion on this measure stops and the Committee will move on to the next measure. If a measure is consensus not reached on a must pass criterion, discussion on the measure will continue, but the measure will not be voted on for overall suitability at this time. And next slide.

For the Committee to vote on a measure, a quorum of 66 percent of active Committee members must be present at the meeting. Since this Committee has one inactive member this cycle, this means that 16 of 24 members must be present at today's meeting for live voting to occur. We currently have 20, so we expect live voting throughout this meeting. For a measure to pass a criterion, greater than 60 percent of the quorum must vote yes. Yes votes are the total of high and moderate votes. If yes votes fall between 40 and 60 percent, then we will have not reached consensus. If a measure is

CNR -- or consensus not reached -- it will move forward to public comment. The Committee will then revote on the CNR measure during the post-comment call. A measure does not pass or is not recommended if less than 40 percent of the quorum votes yes. Next slide, please.

Again, we ask that you do your best to remain in today's meeting for its entirety, but please use the phone line or the web chat to let NQF staff know if you need to exit the call early. The Committee will proceed with discussion and voting as long as we still have quorum. If we lose quorum at any point during the meeting, discussion will continue, but live voting will stop and staff will send a survey link to complete voting asynchronously. Next slide.

Are there any questions about this evaluation process? Okay, hearing none. Next slide. And now I will turn it over to Isaac to do a voting test.

Voting Test

Mr. Sakyi: Thanks, Tami. As stated earlier, you should have received the voting link to the voting platform.

So, on your screen should be a test question, do you like Brussels sprouts. The option is A for yes and B for no. So, we are expecting exactly 20 votes in for this particular vote just to make sure everyone is able to vote. So please let us know if you have any technical difficulties. Please refrain from posting the link to the voting platform in the chat box. It is only for the Standing Committee members.

Member Sood: I don't know about others, but I'm seeing just different screens, severe sepsis and septic shock management bundle screen.

Member Bundy: Geeta, that's what I see too as

well.

Mr. Sakyi: There should be a next.

Co-Chair Septimus: That's exactly what I'm seeing.

Member Yealy: Go to next and you'll see the question.

Co-Chair Septimus: We have to go past this next screen? So start survey?

Mr. Pickering: Yes.

Mr. Sakyi: Yes.

Co-Chair Septimus: Okay, now I see it. Okay, thank you.

Mr. Pickering: So as Tami had mentioned -- and this is Matt from NQF -- Tami had mentioned the platform Poll Everywhere has been updated and so there are some new features that you'll see on your screen. Obviously, that sepsis screen was the first screen you saw. You are able to scroll through to the other questions. However, we kindly ask you to stay on the question in which we are voting on, on your screen, on your end.

So in this case, we are on the do you like Brussels sprouts, and we'll see how far we can divide the Committee on this one. But if you are -- you can scroll to other questions on your end, it's just a new feature of this update on the platform. We kindly ask that you not do that, you wait until we get to the Poll Everywhere question for the criterion at hand and answering the question once it is active, as Isaac will instruct. We do have 18 votes; however, I think we need two more, so ---

Co-Chair Thraen: This is Iona. I'm sorry, this is Iona.

Mr. Pickering: Yeah.

Co-Chair Thraen: My e-mail screen has frozen so I have to restart it. I have not voted.

Mr. Pickering: Okay. Well, Iona, would you like to say yes, you like Brussels sprouts, or no? You're put on the spot now.

Co-Chair Thraen: It depends. I'll say yes.

(Laughter.)

Mr. Pickering: Okay. Okay. So if that was --

(Simultaneous speaking.)

Mr. Pickering: Sorry? Was that Ed?

Member Jackson: Will we need to scroll through the questions or will the question that we are on automatically pop up as the question we're facing?

Mr. Sakyi: You will have to scroll through as we move along.

Member Jackson: Okay, thank you.

Mr. Pickering: Yes, thank you for the question. So this is a little bit different so, again, please, you know if you're having some issues or we have issues on our end, we may ask you to do some refresh, but also to make sure that you're on the right question. It's a new feature of this update of the Poll Everywhere, so we thank you for your patience as we go through this. But, again, you're able to scroll through, but just keep in mind where we are.

Isaac will do a refresher for everyone as we go to voting of what question we're on so that everyone can get oriented, but please do not proceed until we get to those questions. So, we're still on this first

question and I believe with Iona's response, we have 19.

Co-Chair Thraen: I just added mine. I just did mine.

Mr. Pickering: Okay so then we are still missing one person. Was one person not able to vote for the do you like Brussels --- there we go, 20. Thank you. All right, Isaac. Wow, all right.

Mr. Sakyi: Voting is now closed for the question on Brussels sprouts and we have 85 percent of the Committee saying yes and 15 percent saying no.

Measures Under Review

Mr. Pickering: Great. All right, so that's excellent. Thank you very much for voting on that and going through that process and, again, this update that has happened with the Poll Everywhere software, so we'll continue to provide some instruction as we go through today, but I believe we're done with the voting test. It looks like everybody's good to go.

We will now proceed to the next agenda item. I'll turn it back over to Tami, who will just talk about the -- or at least orient us all to the measures under review.

Ms. Funk: Okay, thanks, Matt and Isaac. So some measures are considered complex or fall into categories requiring specific methodological review. The NQF Scientific Methods Panel consists of individuals with methodologic expertise who help review these types of complex measures. Certain measures that do not pass reliability and/or validity are still eligible to be pulled by the Standing Committee for discussion and revote. Next slide, please.

So, this is in fact, 2021, so this is the Spring 2021 Cycle Set of Measures. There are four maintenance

measures for Committee review. Measure 0500: Severe Sepsis and Septic Shock: Management Bundle; Measure 0674: Percent of Residents Experiencing One or More Falls with Major Injury (Long Stay); Measure 0679: Percent of High-Risk Residents with Pressure Ulcers (Long Stay); Measure 3389: Concurrent Use of Opioids and Benzodiazepines. There are two new measures for Committee review; Measure 3501e: Hospital Harm - Opioid-Related Adverse Events; and Measure 3621: Composite Weighted Average for Three Computed Tomography (or CT) Exam Types. Next slide.

The NQF Scientific Methods Panel or SMP independently evaluated five of these measures; 0500, 0674, 0679, 3501e, and 3621. Next slide.

The SMP passed four of the five measures on scientific acceptability. They did not reach consensus on one of the five measures, Measure 3621, Composite Weighted Average for Three CT Exam Types, was consensus not reached on validity. Scientific acceptability is a must pass criteria. The SMP did not reach consensus due to questions regarding level of analysis. The Committee will discuss this as part of their discussions and analysis during the review of Measure 3621. Next slide.

And now I will turn it back to Matt to kick off our discussions.

Mr. Pickering: Great, thanks very much, Tami. And before we go into our first measure as you see listed here, I just wanted to just pause a little bit, leave some room for questions, but before I go there I just wanted to make sure everyone is aware of the process that will happen. So it's been a little while since we all met, right, and this process is good to be reminded of. Each one of the measures has a co-chair that will facilitate the dialogue between the Committee and also if there are

questions to the developer, we can triage those to the developer as needed.

Iona is going to be the co-chair for this measure, but the co-chair will introduce the measure and allow the developer about three to five minutes to provide an overview of the measure. So that could be a description, it could be talking about its importance or addressing any potential concerns related to the measure that they have identified as well through previous evaluations.

After the developer has presented the overview of the measure, Iona will then start the discussion and ask the lead discussants for the measures to present the measure. We will go through each criterion individually and vote on each criterion or sub-criteria first. So we won't be discussing the entire measure up front. So, for instance, evidence is the first sub-criterion under importance to measure and report. The lead discussant will discuss the measure, allow other discussants to discuss the measure as it relates to evidence, and allow the Standing Committee to then discuss and ask questions that will either be triaged to NQF about process or triaged to the developer, if needed, around content. And then once the Committee is done discussing and deliberating evidence, we will then vote on evidence.

After that, we will proceed if it passes on evidence to the next sub-criterion, which is performance gap and then the same process occurs again. The lead discussant presents the performance gap, noting the pre-evaluation comments and also noting public comments that have come in, which is all within the PA, allowing discussants as well to chime in, opening up to the Standing Committee to discuss issues related to the performance gap and then triaging questions accordingly. Then, we vote and then we move on to the next and the next and the

next.

I will just mention as well, we have two composite measures as you've seen previously, 0500 is the first as well as 3621. There are three questions under the importance to measure and report for composites and three questions for validity for composite measures, just keep that in mind, but we'll get to those when we do.

Do we have any questions? I'll pause now to see if the Committee has any questions on that process before we proceed.

Okay. So, hearing none, I will go ahead and turn it over to Iona. Just a reminder that this measure has three recusals. So, Ed Pollak, Ed Septimus and Don Yealy will be recused from this measure, that means from the discussions and also from the voting. The quorum number to achieve is 14, but even with those recusals we have still met quorum so we are good to go with voting for this measure, but we will go back to that once we to the voting. Iona?

Consideration of Candidate Measures

0500: Severe Sepsis and Septic Shock: Management Bundle

Co-Chair Thraen: Thanks, Matt. So this is measurement 0500, Severe Sepsis and Septic Shock: Management Bundle. The measure steward and developer is Henry Ford Hospital and are they on the line?

Dr. Rivers: Yes, this is Emanuel Rivers.

Co-Chair Thraen: All right, Manny, would you like to introduce your measure?

Dr. Rivers: Yes, Sean Townsend --

Dr. Townsend: Yes, this is Sean Townsend. I'm here and I can do that.

Member Fairbanks: I'm having a little trouble hearing. I don't know if that's true for others.

Dr. Townsend: Let me try to change something, then. Perhaps this is better?

Co-Chair Thraen: Yes, I think so.

Member Fairbanks: Yeah.

Dr. Townsend: Okay, very good. Forgive me for a moment, you're running a little early so I had expected to be at my desk, instead I'm on a cell phone heading to my office. I'm going to pull over and I will read our opening statement.

Okay, now we're safe. There'll be no accidents or injuries today.

I'm Sean Townsend. I represent Henry Ford Hospital and CMS as one of the measure stewards for the early management bundle, SEP-1, otherwise known as NQF 0500. I was graciously asked to be one of the measure stewards by Dr. Emanuel Rivers, the key measure developer and world renowned authority on resuscitation and management of sepsis patients. This measure, which I will refer to from here on in as SEP-1, is a composite measure that calculates a hospital's compliance on an all or nothing basis with seven key steps to its resuscitation approaches.

In the prelude to a national sepsis measure, hospitalization rates for sepsis more than doubled from 2000 to 2008. During this time, sepsis was present in greater than 50 percent of U.S. hospital deaths and was the costliest disease at 24 billion dollars annually. Congress encouraged the Centers for Medicare and Medicaid Services to improve

sepsis care after gripping testimony from Ciaran Staunton, father of Rory Staunton, a boy who died of sepsis. CMS identified sepsis as a priority and in 2013, as the measure's application partnership, a multi-stakeholder group convened by the National Quality Forum and required by Congress to review quality measures for CMS, to review measures for possible inclusion in Medicare's Inpatient Quality Reporting Program.

In 2014, the National Quality Forum re-endorsed the measure developed by Henry Ford Hospital, which CMS incorporated as SEP-1 into the Inpatient Quality Reporting Program October 1st, 2015. The measure was officially re-endorsed in 2017 using the 2016 Surviving Sepsis Campaign guidelines as the best summary of the evidence available to support a sepsis measure. These guidelines remain in force today pending publication of updated guidelines.

In 2017, SEP-1 passed the NQF's Infectious Disease Standing Committee prior to our move to the Patient Safety Committee with voting suggesting evidence supporting the measure was ranked as follows: high, four votes; moderate, nine votes; low, zero votes; not voting, one. Voting for scientific acceptability of the measure was similar for reliability and validity. The measure passed CSAC 12 to zero and was re-endorsed. Since these results, no further publications have altered the evidence basis for the initial resuscitation of sepsis patients.

Indeed, in 2016, Surviving Sepsis Campaign guidelines were published. The quantity, quality, consistency of the reports supporting sepsis quality improvement efforts, such as SEP-1, have only increased with the release of results from the New York State initiative, requiring evidence and forum protocol at hospitals throughout the state in the

New England Journal of Medicine. An additional four observational trials have been published supporting timely antibiotics for patients with both severe sepsis and septic shock, noting that hourly delays in therapy are associated with deaths. More to the point, no well done observational or other studies demonstrating deleterious effects of SEP-1 have been published. In this regard, we come before the Patient Safety Committee with only stronger evidence than the NQF Infectious Diseases Standing Committee previously evaluated.

SEP-1 is a percentage of compliant cases quarterly -- rather, is reported as a percentage of compliant cases quarterly at the hospital level based either upon a predesigned random sample of patients or the entire denominator of patients. The measure applies to patients who are admitted to a single facility and transfers are excluded. Recent testing of the measure was assessed at the facility level using CMS clinical data warehouse data since our previous re-endorsement in 2017. This data was presented to the Scientific Methods Panel to assess reliability and validity of the measure for this route of endorsement. The Scientific Methods Panel has already voted to pass SEP-1 this cycle with both elements of the measure development process.

I would like to take a moment to address how we work to incorporate the concerns of providers and hospitals, users of the measure. In November 2019, we were approached by the Infectious Diseases Society of America and other societies regarding concerns they would later publish in a physician paper. In brief summary, IDSA was concerned that the measure needed no reliable start times and that it may promote antibiotic usage in hospitals that was unnecessary. Over multiple subsequent meetings, although we disagreed on these items, we agreed to study a change that would establish a

more reliable start time and prevent the measure from incentivizing the use of antibiotics. This study is now IRB-approved and ongoing and if findings are favorable, we will make these changes.

Separately, CMS became aware that many users had concerns with the requirement that septic shock patients automatically receive 30 mL/kg of crystalloid fluid administration without consideration for the patient in whom this amount of fluid could be deleterious. Although the balance of evidence favors 30 mL/kg, Dr. Rivers and I set about to alter the measure so that providers may select an alternate volume to administer, document it, actually administer that amount of fluid and then pass the measure. This change will become effective in two stages. This July, providers will be able to take advantage of this change if they document advanced organ failure as a purpose or reason. Starting in January, a provider will only need to document the 30 mL/kg may be detrimental.

These changes preempt many of the concerns raised in some letters submitted to this panel. Those letters were no doubt unaware of these changes. We point them out here to underline our commitment to serving the needs of patients, providers and hospitals by adopting changes that reflect the current state of the evidence supporting the sepsis measure.

Since sepsis remains a diagnosis with extraordinarily high mortality, higher than virtually any other common inpatient condition save intracerebral hemorrhage, we must continue to support a national sepsis measure to drive hospitals to innovate quality improvement approaches across the country. We are pleased to present the measure today for re-endorsement. This concludes our opening remarks and we look forward to questions from the Committee.

Co-Chair Thraen: Thank you and thank you for the updated view in terms of incorporating some of the feedback that you've gotten. With that, I'm going to turn it over to Terry Fairbanks, who is the lead discussant, and his team, Elissa, Bret and Jamie. Terry?

Member Fairbanks: Thank you. How's my sound?

Co-Chair Thraen: Your sound's good.

Member Fairbanks: Okay. And I don't want to be redundant, but I thought is it useful for me to just explicitly state the measure? I think Dr. Townsend had it in his description, but unless you say otherwise, I think I should just state it explicitly.

Co-Chair Thraen: Sure.

Member Fairbanks: Okay. So the measure focuses on adults 18 years and older who have a diagnosis of either severe sepsis or septic shock. And consistent with the Surviving Sepsis Campaign guidelines it assesses measurements of lactate, obtaining blood cultures, administering broad spectrum antibiotics, giving adequate fluid resuscitation, vasopressor administration and reassessment of volume status and tissue perfusion, which can include repeat lactate measurement. The data elements and their definitions, I think, we've already discussed. The first three interventions should occur within three hours of presentation while the others are expected to occur within six hours. It is a composite measure, I think we know, it's not risk adjusted.

I do have one clarification, if I can, and then I think I should go into reading the overall summary of the comments. There are extensive comments and there are very varying opinions on this that were submitted and I want to make sure those are well-

represented. But I just want to -- just a note of clarification, with the changes that Dr. Townsend reflected, such as the fluid resuscitation exclusions, are those being considered today, because I hadn't -- I was not aware of those when we were looking at the comments and evaluations.

Co-Chair Thraen: So I think that's a question for Matt. So this is a timing question in terms of new information. I think the review process was prior to that new information. How do we incorporate that into that discussion?

Mr. Pickering: That's correct. So, you are to evaluate the measure that has been submitted in front of you through what the developer submitted and what has been included in the PA, what you have in front of you. If the developer has mentioned that ongoing work is still happening, you may take that into consideration; however, noting that various -- our criteria are to be applied to the measure that has been submitted in front of you, the information to be submitted in front of you and for you to apply to the criteria in that case.

Co-Chair Thraen: So I think -- so Terry, to answer your question about the comments, there's a couple of questions in the chat about questions of the developer, questions from the developer or to the developer can be clarification questions only and then once we start to delve into the endorsement criteria, the questions relevant to those specific endorsement issues can be asked. I don't remember who it was that was asking whether or not you had questions for the developer. Was that you, Geeta? Yes. Please go ahead.

Member Sood: Yes. Just so that I'm clear, clarification questions are okay now and then specific questions related to criteria we would ask later, right?

Co-Chair Thraen: Correct.

Member Sood: Okay. So, first of all, thank you. Clearly we know that sepsis is an important area to focus on and clearly the developers are working on addressing some of the concerns that have been brought up by other organizations, so thank you for that.

My specific clarification question is related to the studies that you mentioned in terms of showing hourly improvement of interventions. Two of the bigger concerns, of course, are that sepsis is a broad category and there's a lot of variability in what exactly sepsis means and that it's a composite measure. So, my question is the data that you are presenting -- that you mentioned, does that address all of the composite aspects of the SEP-1 metric or the sepsis bundle in showing a reduction in mortality or is it specific interventions? I thought it was the latter, but I could be wrong.

Mr. Pickering: And sorry, maybe if I could just chime in as well, it does sound like we're starting to get into maybe a little bit more of the evidence discussion, Geeta, with that question, which is great, that's the first criterion you should be going through. And I think maybe we could start there if we wanted to maybe have Terry present a little bit more about the evidence, what has been presented by the developer. Terry, I agree with you, I think it's great to note the comments of both sides from the public around this measure, so I think that would be welcome there and maybe we can start to open it up for some clarification. Iona, what do you think?

Co-Chair Thraen: Sounds good to me. So, Terry, you want to go ahead and start with the evidence for importance?

Member Fairbanks: Okay, let me just pull that up. Okay, so the evidence's must pass criteria that the evidence here demonstrates that there is some variation in the diagnostic reference levels. Systematic reviews have been provided in the initial piece. There is some -- hold on, I just pulled up the wrong -- you're having me go a different order than my notes were so let me just take a second and get to those.

Co-Chair Thraen: Sure.

Member Fairbanks: Okay, thank you. Sorry about that.

Co-Chair Thraen: No worries.

Member Fairbanks: So the evidence, I think, is probably the most important place to start because that's where if we look at public comments and the data that's been available in the last few years since the last evaluation, that's where I think the questions are raised. I want to endorse what we just heard. This measure is probably, I think, one of the most effective measures we've seen as it came out from the years that it's been out in terms of raising this critical issue for early sepsis, it's made a huge difference for many patients. So I want to really endorse the fact that from the early evidence in Dr. Rivers' study, we really knew that there needed to be much closer attention to sepsis, early treatment and recognition has been critical.

As we get into that, the discussion around the evidence really comes into the fact that this is a composite measure and there is variable evidence in each of the components of the composite, which I think is where the discussion now comes to. Dr. Townsend, I think, did really bring out probably the three biggest issues that are of question here that I think are raising the biggest controversy in

question.

So the administration of crystalloid fluids for both hypotension and the lactate is an area that there is some question about the quality of the evidence. If you look at the three significant studies that have come out in the past few years, looking at these protocols as they've been implemented, there's pretty clear evidence to support early antibiotics and early recognition as being an issue, but the evidence is not as strong when you look at the administration of aggressive fluids.

And I think that was raised well by Dr. Townsend and it sounds like there's a plan to look at that but the question is raised -- and I think we'll get into this when we get into the public comments, some of the society statements and other public comments raise the issue about whether we're causing harm to patients who have cardiac dysfunction, for example, in terms of the aggressiveness of the fluid administration without the ability to make a distinction between hypotensive septic shock and the presence of sepsis. So that I think is probably the largest evidence area. What I'd like to do is just open it up to my other discussants to add, if the Chair thinks it's the right order of business, for us to focus on evidence for a little bit before we move on to the other sections.

Co-Chair Thraen: Yeah, Terry, that's great. So do -- - and either Elissa, Bret, or Jamie, would you like to make any comments?

Member Charbonneau: It was a really good summary. I think that just clarifying that the changes that the developer has suggested that would be implemented do not apply to our current voting today.

Co-Chair Thraen: Correct.

Member Charbonneau: Thank you.

Member Roney: This is Jamie, I would like to add I think although fluids has been debatable since the beginning of SEP-1, you can find empirical evidence either direction and in my opinion, upon review of peer reviewed literature there is a lot more evidence to support the fluid administration and no harm to patients, specifically a few manuscripts that were published within the last couple of years even demonstrating benefit. In one study no harm to those with congestive heart failure and renal disease and liver disease. So I guess that the literature on the fluid is all over the board as we all know, if you live in the world of sepsis. And so I think that fluid could be debated for hours and days and weeks and years and it will be, but as far as everything else, I think although lactate is not specific to sepsis, I think the evidence is very clear on elevated lactate and the impact on mortality, no matter what the etiology of the lactic acidosis is. So that's just my comments on the evidence.

Co-Chair Thraen: Thank you, Jamie. Bret, did you have any comments?

Member Jackson: No, this is still very much a steep learning curve for me. And everything I read, including the public comments, was very supportive of the measure. And so I just appreciated the thoroughness of the work that the NQF staff does to bring this all together. Thank you.

Co-Chair Thraen: Thanks. All right. I'll open it up to the committee as a whole. Is there any specific questions related to the evidence? Go ahead, Geeta.

Member Sood: Sorry. So I just wanted to clarify, because I -- this time, since I'm presenting, I looked very carefully about what exactly each of these criteria are supposed to be. For evidence,

really we are looking to make sure that there's a relationship between the outcome and the health structure and process.

So, if I understand correctly, for evidence, we're not looking for general ideas of whether we think antibiotics are important, but that the entire SEP-1 composite metric influences mortality, which is the outcome or the defined patient population, not just the infected subset of sepsis, but everybody that is coded as sepsis. Is that correct?

Co-Chair Thraen: That seems to be logical. Somebody else is going to comment.

Mr. Pickering: Yeah, I'll just chime in from NQF. Geeta, thank you for doing all of that review of the criteria, as well.

So, just to clarify, this is a composite measure of process measures, right? So, in that sense, we're not looking at tying it to a structure. We're also looking at the association to some sort of desired health outcome, right, the process being tied to some sort of desired health outcome through the evidence. And so the developer would articulate this in their logic model, and support this as well with the evidence they've submitted.

In addition, it's not just looking at the overall composite; it's looking at each one of those components to see if there is strong evidence to support those components of the composite. Also, with that, we have to make sure that whatever is listed in those components has supporting evidence for the population, as well as any potential thresholds, timeframes, et cetera. So that's where we're looking for evidence to support that.

And so there are guidelines within our criteria to assess that. But there's that certainty level that

comes with it. And in this case, the developer did perform a quality-quantity-consistency evaluation to help guide the standing committee to looking at the evidence based on the grading of the evidence, the quality and the consistency of that evidence, to think about the certainty that the standing committee has related to the evidence that's been submitted.

But there still is your expertise that comes into this, as well as your experience within this clinical area, to weigh in on that. So I just wanted to chime in there, Geeta. And I hope that was helpful to answer your question. Okay.

Co-Chair Thraen: If you remember, on page 3 of the document, specifically to the evidence, there was a table that looked at each of the components and then the strength of the quality of the evidence. And so, as it's already been pointed out, lactate has a weaker quality of evidence. Blood cultures is identified as best practice. Broad spectrum antibiotics is a strong recommendation. Crystalloid has got a strong recommendation but with low quality of evidence. Vasopressors, strong recommendation. And volume reassessment was a best practice.

So I think that was quite a bit of detail contained in the overall metric description. So I guess the question is, do you think this is an important performance measure, a measure to evaluate in terms of gap? Do we have people dying from sepsis?

Do we have best practice? Is this an evolving set of knowledge as we go through the process of testing and moving forward? Does this measure actually promote that conversation, promote the ability to do research studies, to test and refine this bundle over the course of time. And should we keep it?

So, it's a must-pass measure. Any other discussion about it?

Member Fairbanks: Well, Iona, it's Terry. I just want to raise a question. I think what you said just now is slightly different than what I heard from Matt, and I think it's really important with this measure because this measure clearly -- as I said in the beginning, there are few studies and measures that have done more to draw attention to an area of concern of immediate recognition of a critically dangerous condition, and early treatment of that condition. The numbers, you all know the numbers, but it's one and a half million people or more that suffer from this each year in the U.S.

So it's critically important, which I think is what you were talking about. And it clearly draws attention to it. But when we're looking at evidence here, I think that the question that we are being asked is, since this is a composite measure, is there either evidence as time goes on -- because this is different than the initial -- when this was initially brought in, there was good evidence to show promise. Now we have the benefit of a host of evidence after several years of implementation.

And I think the question becomes: is there evidence to either raise a question about the safety of the unintended consequences of components of the bundle that have less strength -- and I'll go back, but we've mentioned a few -- r is there some question about the overall impact? And there have been at least three studies in major journals, like New England Journal of Medicine, that do raise some of those questions. And so I think --

Co-Chair Thraen: So, Terry, just for clarification, you said they raise those questions. Are they providing evidence of harm that raises those questions? Or are they philosophically raising those

questions in terms of the evidence?

Member Fairbanks: I think the important thing is that they are not showing evidence of benefit. And I think it's critically important that when we're looking at a measure we have to have a bundle that we know works. We need it to be a simple bundle. We need the evidence -- we have the ability right now to re-evaluate over time and look at the data.

Dr. Townsend pointed out two significant issues that I think are probably the most important. Fluid resuscitation is probably the most important of those, but also the potential for increasing unnecessary antibiotic use.

But the other pieces that were mentioned that are just as important is that the measure and remeasure of lactate is weak. I'm sorry, the blood cultures is a best practice statement, but it's still weak.

And so, the question is, as time goes on, we have to hold ourselves to a higher standard. Is there evidence and data to support that each of these components is what we thought it would be? And if the answer is that we're not sure of that, then we have to question whether there's a possibility that there could be harm.

And I really like what Dr. Townsend said about what they're going to study. But I'm concerned right now that we don't have that information. So the question is: do we want to continue on with a metric that they're going to study two areas of potential lack of impact or potential harm?

Member Sood: I would just echo that, and I think it's really important. So I appreciate, Terry, the way that you said that and your eloquence in describing and framing the question, because I don't think the

question is: is this an important area, is this going to foster improvement? I think the question, like Terry said so beautifully, and correct me if I'm wrong, is: is there evidence that this specific metric does not cause harm and does specifically improve the outcomes that we want it to improve? At least, that's how I'm understanding the evidence. I think you expressed that very well.

Co-Chair Thraen: Are there any other different comments to be made prior to voting?

Member Roney: This is Jamie.

Member James: I'd like to ask a question. So I'm by no means an expert on sepsis, at all. But I look at the weight of evidence. And those of you who are knowledgeable in this area say, unequivocally, the weight of evidence is that this thing works to the benefit of patients. When risk and benefit are compared, what is the weight of evidence in favor of this?

Member Sood: Terry, feel free to weigh in. But I'm not sure that's completely clear in all of these circumstances, because the weight of evidence is partly on the patients that you're studying. There's excellent evidence that if somebody is truly infected every hour of antibiotics will make a difference. When you start parsing out the different pieces and the different patient groups, my perception is the evidence is less clear. But I would love to hear from the other clinical team members to see what they think of the evidence party.

Mr. Pickering: And before we go there -- sorry, Terry. Didn't mean to cut you off. I do see that other people from the public are contributing to the chat. I will just mention that there is a public comment portion later on in the meeting. So if you have comments you'd like to share in relation to

anything that's discussed today, please reserve those for the public comment portion. So that means refraining from putting anything in the chat, which the chat would be reserved for our standing committee. Thanks. Sorry, Terry.

Member Fairbanks: Thanks.

Dr. Townsend: May I ask a question?

Member Fairbanks: I just --

Dr. Townsend: Could I ask a question for just a moment, if you don't mind?

Co-Chair Thraen: Go ahead.

Dr. Townsend: It's Sean Townsend. Is it appropriate for the developer to chime in in this discussion at any point? Or is it amongst the panel only?

Co-Chair Thraen: Unless there's a point of clarification that you want to give information about. So, for example, do you have evidence of harm as a result of this measure that you could bring to the table?

Dr. Townsend: No, I just wanted to -- what I was hoping to discuss was the evidence we provided, which I appreciate the discussion so far, but it hasn't really touched on the evidence in the packet that we submitted showing the mortality reduction in our analysis, plus the several recent studies of sepsis bundle care with the individual elements that show reductions in mortality. And so if it's possible, I'd like to remark on those.

Co-Chair Thraen: Okay. Hold on a minute. I think Curtis was going to ask a question first, and then we'll get back to you.

Member Collins: Yeah, I guess my question was for

Dr. Townsend, and probably what you'll discuss here. In terms of both the IDSA comment, which I know Ed was part of, as well as some of the evidence there, I'm really interested in your comments about the benefit, particularly in the severe sepsis versus septic shock population for early antibiotics. I know you said you were going to study that moving forward.

However, there's an excellent study by one of our pharmacy colleagues, Amy Pakyz, that did show increases in antibiotic use. I've seen some of your commentary pieces, which thank you for contributing those, but I would ask if you could kind of discuss some of the comments, particularly the ones by IDSA regarding severe sepsis and the potential for increased antibiotic use, with or without some benefit there.

Co-Chair Thraen: So you're talking about the risk? Okay. Any other questions from any members of the committee?

Mr. Pickering: I see Jamie has her hand raised. And, Jamie, I'm not sure if that's from previously. And also Jason had a comment within the chat box. I don't know.

Co-Chair Thraen: Thanks. Jamie, did you have another comment?

Member Roney: Yes, ma'am. I just wanted to raise my hand because, you know, just to be in the queue. I just wanted to say I know antibiotics stewardship is important, and I think our recent COVID-19 pandemic has showed us what could happen if we didn't have antibiotics available to give people with infections across the world today.

But I do think with COVID-19, if you look at NIH's report of coinfection, more than half of our group of

people that had COVID-19 infection had a coinfection. And there was just a publication of causes and outcomes of sepsis in Southeast Asia that demonstrated anywhere from 7 to 12 percent of viral infections had a co-bacterial infection.

And so although I appreciate the concerns about early antibiotics and antibiotic stewardship, as long as we de-escalate that, I still see more value, based on Kumar's work that's been supported through several other studies on early antimicrobial therapy, that until we can target the organisms that are leading to what we're looking at, that early de-escalation -- it's just still hard for me to weigh there's more harm than benefit to early antibiotic administration and suspected severe sepsis, septic shock.

And that's just my perspective. But I know coinfections are large. And just the risk of the microbial not being targeted and not being eliminated earlier than later is well established in the literature. And so I see both sides of the argument. It's like fluid.

But I do feel like, also, some of these study designs seem to be flawed in themselves in the design to try to make the argument to one side or the other, especially around fluid administration. And so it's very hard to take the bias out of such a hot topic like sepsis. And so to look at the empirical evidence through an unbiased lens, it's really difficult, I think.

Co-Chair Thraen: Okay.

Member Fairbanks: Can I ask a point of clarification just around that comment, so that we can focus the discussion? Because I think we're going to vote first on evidence. And I came into this, as I was preparing, thinking about the evidence to support the components in the composite. There's no doubt

about the critical importance of sepsis. But the comment just now questioned if we're looking at whether the good outweighs the bad. And that's not how I was approaching it.

So can we have a clarification on that? Because -- and then while I have the floor, I'll just say, I think it's important, before we vote, that I do summarize the public comments, because we have several societies and we have kind of opposing and for. And they talk a lot about evidence. That's kind of the main theme in it. So I have that clarification and that suggestion in the order of things.

Co-Chair Thraen: All right. So I'm just going to read to you the definition, and then we'll follow up with the other comment.

So, the importance to measure and report, extent to which the measure focus is evidence-based, and important to making significant gains in healthcare quality where there is variation in, or overall less than optimal, performance.

And then you wanted to summarize the public comment. And I believe the developer also wanted to summarize the impact on mortality from the data that they had.

So, go ahead, Terry. Start with the comments, and then we'll go back to the developer. And then we're going to take a vote after that.

Member Fairbanks: Okay. And procedurally, I just want to get your endorsement. This is going to take a few minutes, I think. But I feel like it's the right time to go through that. Is that okay from a procedural standpoint?

Co-Chair Thraen: Matt, are you okay with that?

Mr. Pickering: Yeah, I think it's good, too. I do want

to come back to some of the -- like, Jason had some comments in the chat box. And then I know the developer, we were going to circle back to the developer for their response.

Co-Chair Thraen: Okay.

Mr. Pickering: But I think it's good to recognize the public comments, because they were quite varied. So, Terry, I'll turn it to you, and maybe we can come back to some more of the Q&A and developer responses.

Member Fairbanks: Okay. And I'm going to do these in the order that they appear in the document. And they're somewhat varied in their position.

So, the first is the American Medical Association, which gave a statement. They first said -- and I think this reflects probably everyone, all of us here -- that they support the intent of this. I think there's no question in anyone's mind how critically important a measure around severe sepsis and septic shock is.

They said that they do not agree that this composite measure meets the need to do this. And, therefore, they're urging us to recommend removal of the endorsement until there's more alignment from more studies. They felt that there was negative unintended consequences such as incentivizing antibiotic use. They do cite the IDSA position paper, but they have a separate statement so I'll skip that.

Also, besides antibiotic overuse, they were concerned about what was mentioned earlier, the case of patients with severe systolic dysfunction, heart failure, and the negative impact it can have on the patients. And, of course, our role is supporting all patients and keeping them safe.

So, that is something that I think Dr. Townsend

addressed early on. It sounds like they have a plan to look at a way to correct that. I think that -- and that's my clarification, is whether we should wait and reevaluate the metric once that plan is in place? Because right now we're evaluating without that.

So then the second piece is the Society of Critical Care Medicine. They support continued endorsement of NQF. They cite the major public health concern. They cite, between 2008 and 2014, feeling the measure -- they cite that the measures were comprehensively reviewed and vetted by multiple stakeholder groups. And they felt that they would -- their final statement, which I think is important, is that they endorse the ongoing process of NQF maintenance to bring the measures into alignment with the latest published evidence as a stimulant to implement evidence-based practice.

The Sepsis Alliance cites -- they cite how large the sepsis issue is and how many millions of people are affected. And they cited the fact that the SEP-1 focuses on timely recognition of sepsis and early intervention. And they do support us continuing. They specifically said they disagree with those who urge removal of the measure, and saying they're embracing the dual responsibility to diagnose and treat sepsis in a timely way and manage our antimicrobial -- the antibiotic overuse issue. They also recognize the issue with the judicious use of IV fluids as well in their statement.

There's a comment by Bruce Quinn saying that he -- his comment is driven by a systematic review of the published evidence, and that he feels it's fundamentally different today than it was in the performance of the randomized controlled trial intervention initially. He feels it is no longer necessary to make decisions based only on RCTs. Rather, we have a direct volume of evidence as how well a measure of performance and correlated with

real world patient outcomes. And that he feels that correlation is not very strong. He goes on to cite some examples of hospitals which are not correlated.

The New Jersey Hospital Association expresses strong support of the severe sepsis and septic shock bundle, the SEP-1. And they cite, in a letter to the editor in JAMA, in 2016, CMS voiced rationale. And they cite several other reasons that, I've stated some already, of being in support.

The Society to Improve Diagnosis in Medicine supports continued endorsement of the measure. And they cite the inaccurate or delayed diagnosis as the most common issue that causes -- that goes on to catastrophic events for patients and their families. They did note they recognized the current sepsis measure is imperfect and needs to be updated, I think as we've already mentioned in different ways. They strongly urge that NQF work aggressively to update the measure based on the latest evidence. And they're really emphasizing the importance of having a measure around sepsis.

The American College of Emergency Physicians again stated the importance of sepsis and the importance of early recognition. But they felt that they would not be in support of continuing. They believe the measure would be markedly revised to be continued.

And I should say -- this is not a conflict of interest, but I think important to recognize -- I'm an emergency physician. This is my professional society, but I did not have any role at all in the development of their position statement. So I'll just make that statement.

So, they support the recommended revisions as we've heard about today, and as outlined in the

Clinical Infectious Diseases article that was noted earlier.

Leapfrog Group strongly supports continuation of the SEP-1. And they really emphasize, as we've talked about, the critical importance of the early diagnosis and treatment of sepsis.

And then, finally, there were five NQF members who submitted. Two supported the measure and three did not support the measure.

So, thank you for allowing me to do that. So many of these comments that came in were around evidence. I felt like it was an important order to be able to have everyone see those and have their attention drawn to them, all the public input, before voting.

Co-Chair Thraen: Thanks, Terry. Jason, did you want to make a comment?

Member Falvey: Thank you. I mean, I commented a little bit on the evidence that was updated from the developers in 2018, a published study that was done in New York state. So, implementing these mandatory reporting of sepsis bundle, very similar to this measure that we're discussing. And they presented some pretty clear mortality benefits per element of the bundle.

So I think that's helpful evidence to look at. Some were certainly stronger than others. But, on a relative scale, these improvements in mortality were 10 to 20 percent difference. So it's not a small magnitude of benefit in a relative sense.

And I also think a focus on mortality, as we've gone through with COVID, probably misses the fact that poor management of sepsis early leads to more likelihood of mechanical ventilation and disability and a lot of morbidity down the road. So we're

probably missing the burden of poor management by just focusing on mortality. So I think the importance of this measure -- and if we're seeing that much benefit in mortality, then we're probably seeing even greater benefit in disability and other symptom burden down the road.

Co-Chair Thraen: All right. Thank you. And then, finally, the developer, you wanted to update on mortality question?

Dr. Townsend: I do. Thank you very much. This is Sean Townsend again. So, I just want to frame our conversation carefully. I want to sort of indicate to the committee that the measure under consideration today is the measure presented, not the one with the future changes.

And we're very specific in doing so because I'll remind the committee that the measure passed in 2017 based upon a systematic review of the 2016 Surviving Sepsis Guidelines. The burden I see in the evidentiary piece is to demonstrate that there have been new studies since that time that demonstrates somehow or another the measure would cause harm or that specific elements could cause harm. And yet today in our conversation -- when we talk about evidence, I think we need to be specific -- I have not heard anyone to date cite a specific study with a specific name showing a specific harm and quantifying it.

And so our discussion I think has been a bit amorphous, and I've been surprised about that since we're about to vote on this measure. I think that the evidence presented in the measure itself, the packet for submission to the committee, is not being (audio interference) --

Co-Chair Thraen: I'm sorry. You're fading out. You're cutting in and out.

Dr. Townsend: Is this a little better?

Co-Chair Thraen: Yes, it is.

Dr. Townsend: Okay. I was going to say that the evidence we presented in the measure submission packet has also not been considered today before the committee. We showed a five percent reduction in mortality in a carefully controlled propensity score analysis. We are very careful of this analysis to demonstrate that there is a reduction in mortality and that it is associated with the individual elements.

This correlates with the evidence since the 2016 guidelines in the New York State study analysis published in the New England Journal of Medicine that shows a consistent mortality reduction with compliance with sepsis bundles. And, again, all the elements in the New York State initiative are the same elements that existed in SEP-1. The only element that was not shown to be significantly correlated independently with a reduction in mortality was the application of vasopressors. And it's well known that that's confounded as a marker of severity of illness.

So, both the SEP-1 bundle and the New York State study share commonalities that are highly likely the same, yet we haven't considered these specific pieces of evidence or the evidence submitted in the evidence packet showing a five percent reduction in mortality.

I'll conclude with remarks on antibiotics. I was asked to specifically address the possibility that the measure increases the use of antibiotics unnecessarily. Again, I've heard one study cited in the course of this discussion, at a single center, that showed that there was an increase in use of antibiotics for the treatment of urinary tract

infection. And that study did not show that there was harm from increasing the use of antibiotics in that single center study. It has not been reproduced. It was not blinded. It was not controlled. So I don't believe that this study is indicative of the entire nation's experience with the SEP-1 measure.

Finally, I'll just note that since this measure was passed in 2015 CDC has reported a decrease in all types of healthcare associated infections on an inpatient basis, including C. diff, extended spectrum beta-lactamase infections, and other resistant organisms. So there's no evidence in the record of harm associated with this measure. And I've heard none specific today.

Co-Chair Thraen: All right. Thank you. So I'd like to lean towards calling for vote.

Mr. Pickering: Sorry.

Co-Chair Thraen: Go ahead.

Mr. Pickering: Yanling has her hand raised. Sorry.

Co-Chair Thraen: Okay. Yeah, Yanling. Sorry I missed you.

Member Yu: No problem. Thank you. I just want to really -- I heard all the concerns, but I really want to present the patient's perspective.

I have a friend whose son died of sepsis, a young son, young boy. And it's because the late intervention. So if we look at the harm, everything has harm and risk. And the way is to look at the balance, whether there's more harm or more risk. For a sepsis situation, there's so many patient died of it, and early intervention is so critical to save lives.

As far as the antibiotic stewardship, it's definitely an excellent point to raise about this. But there's so many other misused antibiotics in other units and situations. I feel sepsis is not the place to really raise this when timing is so critical and people's life is on the line.

And then about fluid resuscitation, and definitely when a patient who had congestive heart failure, you don't want to pump up lots of fluid. But I think if you do carefully, and do titrate, monitor patient carefully, I think this is a manageable situation for patient with the heart failure. So I just want to emphasize the risk-benefit about this. In particular, even the sepsis, so devastating consequences to people's life.

Member Fairbanks: Yeah.

Member Yu: Thank you.

Member Fairbanks: Thank you. I think it's such a great way to wrap up the discussion. The patient, I think we all have to remember we're here to instigate the highest quality care to our patients.

And I really want to emphasize I think this discussion, it's important to remember, this is not a discussion around whether or not sepsis -- early recognition, aggressive treatment of sepsis is important. We know that it is. It's a question about whether this particular measure has ongoing evidence in support.

Again, we have the advantage of a huge of amount of data that we didn't have when we first looked at his measure. And there's a lot of discussion around evidence. I think the burden here is to demonstrate the benefit without any adverse consequences. There are now three major studies, the ProCESS, ARISE, the ProMISe study. The PRISM investigators

did a summary of all these studies together, all of which are raising the question about whether there is benefit to all the components in this.

Dr. Townsend: I would disagree fundamentally with that. And I need to interject here. I disagree fundamentally with that. This is Sean Townsend. The enrollment criteria for those studies were basically the admission criteria for SEP-1. Those patients received fluids, antibiotics, lactate checks, all of that to get into ProCESS, ProMISe, and ARISE. So it's not the case, at all, that those studies raised question about these interventions. Those studies raised questions about early goal-directed therapy.

Member Sood: So, before we go down this road, just to clarify, is the developer supposed to be part of this discussion part, because I --

Co-Chair Thraen: I think that -- excuse me, Geeta. I think that the developer -- this is kind of an open-ended conversation for clarification purposes. And if he's got -- if the developer has a point of view that clarifies that, I think that's appropriate.

But I think all of the issues have been put on the table. So, we're not here to resolve this question. We're here to vote as a group on whether or not we're convinced that there's enough evidence to move this forward.

Member Fairbanks: Right. Iona, I would just submit, I think the developer has a conflict of interest similar to the conflict of interest that some of the other discussants that were stepping out had. So I think we have to be careful about -- I think if we ask for a clarification, it's good. The debate that we just had is a question. I don't think we have the answer to that. And so I think it's important that we separate those two things.

Co-Chair Thraen: So, I'd like to call for the vote. I don't think there's any new information. I think the issues have been fully vetted on both sides. So I'd like to call for the vote.

Mr. Pickering: And is that -- yes, anything new for anyone else just to make sure before we go? I'm just checking the chat. Any new information that needs to be discussed related to evidence?

Again, this is evidence for this composite measure, right, thinking about the association of the processes, the components listed within the desired health outcomes, and each component itself, looking at the evidence to support the components. I don't see anything in the chat or hands raised. Okay. I'll turn it to Isaac. Go ahead, Isaac.

Mr. Sakyi: Thanks for that. Voting is now open for Measure 0500 on evidence. And just a reminder, we are expecting 17 votes.

Mr. Pickering: Right. So that means Don Yealy, Ed Septimus, and Ed Pollak would be removed or recused from this vote. So out of 20, we should expect 17. And again, you'll have to scroll to this question on your end. So it's the Importance to Measure and Report 1a, as Isaac mentioned. I've seen 12 votes, 13.

Member Yu: My voting link still shows a special section.

Co-Chair Thraen: You need to click next at the top right-hand side to get to the next question.

Member Yu: Got it. All right. Thank you.

Mr. Sakyi: Okay. We have all 17 votes in. The voting is now closed for Measure 0500 on the evidence. We have 3 votes for high, 9 votes for moderate, 4 votes for low, and 1 insufficient. With

12 votes between the high and the moderate, the measure passes on evidence.

Co-Chair Thraen: All right. Moving forward, the next section is the scientific acceptability.

Mr. Pickering: Sorry.

Co-Chair Thraen: Sorry. There's -- go ahead.

Mr. Pickering: Yeah, performance gap.

Co-Chair Thraen: Performance gap, you're right. Sorry I missed it. I keep thinking they're one. So performance gap, Terry, do you have any comments about that?

Member Fairbanks: Excuse me. There's a performance gap and disparities. The -- sorry. Hold on. Again, my notes are jumping to the other measure. Sorry.

Co-Chair Thraen: Okay.

Member Fairbanks: This happened last time, and I should've been ready.

Member Roney: You'll be ready for next time, Dr. Fairbanks.

Member Fairbanks: I know. Okay. There were disparities noted in age, gender, and race. So from that, there was a performance gap submitted. The preliminary rating was high, and I can just open up any other comments to my co-discussants.

Member Roney: Terry, I'd like to comment on that. This is Jamie. It's obvious that sepsis is equal opportunity, and there was an incidence and trends of sepsis in U.S. hospitals using clinical versus claims data in 2009 to 2014, published in 2017.

And that shows the disparities. But I was looking at

questions about risk adjustment. And I don't know that it points towards gender. There are at least two landmark studies that show men are at increased risk of sepsis than women. But I don't know that risk adjustment is necessarily as much related as zip code.

I think the disparities we're seeing in outcomes may potentially not be related to at least what we're considering here which is general, probably more zip -- more location and/or comorbidities if they're not part of what we're looking at. So that's just my comments there. I don't know how you could risk adjust it unless you were risk adjusting for comorbidities.

Co-Chair Thraen: Thanks, Jamie. Any other discussant comments?

(No response.)

Co-Chair Thraen: All right. Any comments from the developer about performance gap?

Dr. Townsend: No particular comments. I rest on the evidence we submitted.

Co-Chair Thraen: Okay. Any questions from the committee members on performance gap?

Mr. Pickering: I'll just add too. Just I know we're talking about disparities. And we talked that disparities is part of performance gap, right, to see if this measure is able to identify disparities across sub-populations. And the performance gap is looking at the range around some sort of central tendency with this measure. But it's also considering if it's less than optimal performance, just factoring those three different types of data points into the decision making of performance gap.

Co-Chair Thraen: Okay. All right. Let's call --

(Simultaneous speaking.)

Co-Chair Thraen: Go ahead.

Mr. Pickering: Sorry. I see a hand up. Sorry, Iona. Just came up. So Jason has his hand raised.

Co-Chair Thraen: Jason?

Member Falvey: Hi, it's just a quick question, and maybe the developer can answer this. It looks like all of the disparity testing was done at the individual patient level. And I'd like to know if there's any consideration of doing this at the facility level because there's quite of bit of evidence looking at -- in other nursing home settings and home care settings at facilities that care for a high percentage of minoritized patients or patients on Medicaid may have different outcomes. So if a Black patient is being served at a very high quality, majority white hospital, the disparity might not be as evident as a hospital that is treating a majority of minoritized patients that might be underresourced.

Dr. Townsend: I'd answer that if I'm allowed to.

Co-Chair Thraen: Yes, go ahead.

Dr. Townsend: Thank you. In the analysis that we submitted, the propensity score matched analysis that demonstrates these disparities. We matched the patients. You're right. It's a patient level analysis.

However, it was done with a mixed effects model considering hospital as a factor. So we factored this into our calculations. And so the disparities you see reflect both the patient level and a consideration for the hospital itself, if that directly answers your question. And we do notice disparities in the application of the measure to various races and ethnicities.

Member Falvey: Thanks. No, that's helpful. The random effect piece definitely answers that question, so thank you.

Co-Chair Thraen: Thanks, Jason. Any other questions that we missed?

(Simultaneous speaking.)

Mr. Pickering: I don't see any in the chat box or any other hands raised.

Co-Chair Thraen: Let's call for the vote.

Mr. Sakyi: Voting is now open for Measure 0500 on performance gap. The option is A for high, B for moderate, C for low, and D, insufficient.

Mr. Pickering: And once again, you'll have to scroll to this question on your end.

Member Fairbanks: Is this Question 3?

Member Jackson: Yeah, I think maybe listing the question number will be helpful since we have to scroll to it each time.

Mr. Pickering: Yes. So this is an Importance to Measure and Report 1b, performance gap, 1b. That's what it should say.

Member Fairbanks: Yeah, on our end, there's a question number.

Co-Chair Thraen: Number 3.

Member Fairbanks: Okay.

Mr. Pickering: Question No. 3. Thanks, Iona.

Member Fairbanks: And then should we press next when you tell us that we're on the next question? I feel like I'm managing this interface differently than

it's worked in the past.

Co-Chair Thraen: It keeps flipping back to the previous question. So you have to kind of keep scrolling forward. It doesn't seem to hold.

Member Roney: That's what it's doing to me as well.

Member Fairbanks: Yeah, yeah. So do we have to press the next button in order for you to record the answer?

Mr. Sakyi: No, once you vote, it is recorded.

Member Fairbanks: Yeah? Okay.

Mr. Sakyi: And you hit next when it's time to go to the next question.

Member Fairbanks: Maybe that's what I can --

(Simultaneous speaking.)

Co-Chair Thraen: See if it'll hold.

(Simultaneous speaking.)

Mr. Sakyi: Voting is now closed for Measure 0500 on performance gap. We have 6 votes for high, 9 votes for moderate, 2 votes for low, and zero insufficient. Measure passes on performance gap.

Co-Chair Thraen: All right. Thank you. All right. Now we can go into scientific acceptability, reliability, and validity. Terry, you want to --

Mr. Pickering: Actually --

Co-Chair Thraen: -- start with comments?

Mr. Pickering: -- not yet, not yet. We have the quality construct portion of it. So again, there's three questions for -- or the importance measure

report when it comes to the composite.

Co-Chair Thraen: I'm making notes --

Mr. Pickering: So there's the --

Co-Chair Thraen: -- on my cheat sheet.

Mr. Pickering: Sorry, sorry.

Co-Chair Thraen: No worries.

Mr. Pickering: You're doing great.

Co-Chair Thraen: Quality construct, what does that mean?

Member Fairbanks: It's an all or none measure. It's a composite that the preceding components all have to be met in order to meet the measure. There was no written committee feedback solicited specifically about this that I saw, but there were no concerns raised that I saw unless one of my co-discussants saw something I missed.

Co-Chair Thraen: Any comments from your discussants, Jamie or Bret, on the list?

(No response.)

Co-Chair Thraen: All right. Any questions from the committee?

Mr. Pickering: Jason, I still see your hands raised, so I'm not sure if you have a question. But I'll just sort of -- or Iona, clarify your question before maybe we see if Jason has one instead. The quality construct, this portion of the Importance to Measure report is to see if the construct itself is -- has good rationale and seems logical.

It's logically articulated as well as the description of how it's calculated, so this aggregation or weighting

as Terry has described and what the developer has reported out. So that's what this question is getting at for the Importance to Measure report. But Jason, I don't know if you had your --

Co-Chair Thraen: He said he lowered his hands, so he's good.

Mr. Pickering: Okay, great.

Co-Chair Thraen: And you guys, the survey keeps kicking everybody out and you have to restart the survey and then proceed to that question. So we're on Question 4, 1c, composite.

Mr. Sakyi: Voting is now open for Measure 0500, Question No. 4, composite, quality constructs. The options are A for high, B for moderate, C for low, and D, insufficient.

Member Roney: Matt, I will say the script that you emailed us helps a lot because I'm trying to follow by the script. I'd never seen that before, and it was helpful. So I thought I'd share with you and your staff.

Mr. Pickering: Well, thanks, Jamie. We try. We recognize there's a lot of information, and we try. And everyone is doing a fantastic job so far. So thank you very much. And it looks like we're missing one more vote. There it is.

Mr. Sakyi: Voting is now closed for Measure 0500, the composite quality construct. We have 6 votes for high, 9 votes for moderate, 2 votes for low, and zero insufficient. The measure passes on the composite quality construct.

Co-Chair Thraen: All right. Now can we move to scientific acceptability? I'm gun-shy now.

Mr. Pickering: I'm sorry, Iona. Yes.

Co-Chair Thraen: It's all right. I'm just teasing.

Mr. Pickering: Yes.

Co-Chair Thraen: Okay. Terry?

Member Fairbanks: Okay. Reliability, there were no initial -- no major issues found in the initial rating. The developers showed a reliability of 0.92. There -- I would just defer to my co-discussants if there are any other issues they have.

Member Roney: I have no other issues.

Co-Chair Thraen: Okay. Any questions from the committee members on reliability?

Mr. Pickering: I'll just add as well that this measure was reviewed by the Scientific Methods Panel. So they are again assessing reliability, validity, and there's also a composite question as well. Again, for validity, there's that other question around the composite for the empirical testing of the composite.

But we're talking about reliability. So in the review of reliability, the SMP did pass the measure on reliability with no major concerns as Terry had mentioned some of the reliability results as well. So I just wanted to make that -- note that it did -- it was reviewed by the Scientific Methods Panel and did pass on reliability.

Co-Chair Thraen: Okay. Any questions from the committee members?

Member Sood: I have a question/comment. As I understand, reliability is -- it's the specifications part and the testing part. And for the specifications, it's whether it's consistent and also included in the way that NQF defines it valid.

In other words, is the definitions representing what we want it to represent? So I guess I'd just add that I have some concerns that the claims-based sepsis definition that I'm not sure is completely reliable in terms of assessing the patients that we really want to identify. So I wanted to just bring that up.

Co-Chair Thraen: All right. Thank you. Any other comments or questions from the committee members?

Dr. Townsend: If it's appropriate again, as the developer, I would respond. But if it's not, I'll hold me thoughts.

Co-Chair Thraen: Does any of the committee members want the developer to respond --

Member Roney: This is Jamie.

Co-Chair Thraen: -- to the question?

Member Roney: And I don't mind if Dr. Townsend responds.

Co-Chair Thraen: All right.

(Simultaneous speaking.)

Member Sood: I don't think I'm asking specifically for clarification. But sure, I mean, it's that same concept of I'm not sure where the lines are for conflict of interest, but whatever Iona or Matt think is appropriate.

Co-Chair Thraen: So the question is you've made a comment and to the committee members. Does anybody need clarification on the comment that's been made by Geeta from the developer?

(No response.)

Co-Chair Thraen: All right. Then we're going to

move forward and call for the vote.

Mr. Sakyi: Voting is now open for Measure 0500. And the question is, do you accept the Scientific Method Panel's high rating for reliability? The option is yes and no. Go ahead --

Co-Chair Thraen: And this --

Mr. Sakyi: -- Iona.

Co-Chair Thraen: I'm sorry. This is Question No. 5. So scroll to Question No. 5.

Mr. Pickering: Yeah, thanks, Iona. I appreciate everyone's patience with this new software update. And again, just to remind folks, since it's been reviewed by the SMP and passes the SMP, you are asked if you'd like to uphold that rating. So this is what it's asking.

Co-Chair Thraen: So Matt, this is a true example on it takes a village.

Mr. Pickering: Couldn't do it without everyone here, that's for sure.

Co-Chair Thraen: We could not. We could not.

Mr. Sakyi: Okay. It looks like we have 17 votes in. Voting is now closed for Measure 0500. We have 100 percent voting yes to accept the Scientific Methods Panel's high rating for reliability.

Co-Chair Thraen: All righty. Terry, validity?

Member Fairbanks: Validity also was looked at by the SMP. They found no major issues with validity. They noted the data element validity testing found moderate to high agreement in strong majority of the data for 15 of the 19 elements. And the elements with weaker agreement were more rare in

nature, they noted. So the score level validity testing found a strong inverse relationship between facility mortality rate and measure pass rate.

Co-Chair Thraen: Any comments from the discussants?

(No response.)

Co-Chair Thraen: Any questions from the committee members or comments?

(No response.)

Co-Chair Thraen: All right. I call for the vote.

Mr. Sakyi: It's now open for Measure 0500. The question is, do you accept this Scientific Methods Panel's moderate rating for validity? The options are A for yes and B for no.

Co-Chair Thraen: Question No. 7.

Member Sood: To Question 6, is that skipped, or --

Mr. Pickering: Yes. So --

Mr. Sakyi: It is skipped because we accepted the scientific acceptability rating.

Mr. Pickering: Right. If the standing committee did not accept it, we would have to vote separately on that question. And since the standing committee 100 percent accepted the reliability rating from the SMP, we did not go to that question.

Member Sood: Thank you for teaching me that.

Mr. Pickering: Normally, you wouldn't have to worry about that with the software update. It's different.

Mr. Sakyi: It looks like we have 17 votes in. Voting is now closed for Measure 0500. We have 100

percent voting to accept the Scientific Methods Panel's moderate rating for validity.

Co-Chair Thraen: All righty. Next question, next criteria is feasibility. Terry?

Mr. Pickering: Hey, Iona, no --

Co-Chair Thraen: No?

Mr. Pickering: -- we got to go back. So it's just around the composite construction portion. So this is just any sort of empirical testing with a composite construction.

Co-Chair Thraen: Okay.

Member Sood: And --

Co-Chair Thraen: I lied.

Member Sood: -- Matt and others, could you remind me? I don't remember seeing if that was addressed in the Scientific Methods meetings. Do you -- could somebody remind me about that?

Mr. Pickering: Yeah, so this is just looking to see if the empirical analysis could demonstrate that the component measures really add value to the composite and that the aggregation of those and waiting of the results are consistent with the quality construct. And the Scientific Methods Panel did evaluate that as well and also did pass the measure on that element. So this is another question just asking about the empirical testing of a composite itself. And the Scientific Methods Panel did evaluate and then also passed it as well.

Member Sood: Thank you.

Co-Chair Thraen: Terry, do you want to make any comments about that, composite construction?

Member Fairbanks: There were no issues raised about composite construction. And I'll note that it hasn't changed since the previous ratings, so that's a positive piece.

Co-Chair Thraen: All right. Any comments for the discussants?

(No response.)

Co-Chair Thraen: Any questions from the committee members?

(No response.)

Co-Chair Thraen: All right. Now we can call for the vote. I've gone to my cheat sheet details. So I'm back up to speed, you guys. Apologize. I was using the global definitions.

Mr. Sakyi: Voting is now open for Measure 0500. And the question is, do you accept the Scientific Methods Panel's moderate rating for composite construction? The option is A for yes and B for no.

Member Fairbanks: Question 9, I'll note.

Co-Chair Thraen: Thank you. It's a challenge in multitasking.

Mr. Sakyi: And we have exactly 17 votes. Voting is now closed for Measure 0500. We have 94 percent voting to accept the SMP's rating and 6 percent voting no.

Co-Chair Thraen: Okay. Now feasibility.

(Simultaneous speaking.)

Co-Chair Thraen: Terry? Oh, sorry.

Mr. Pickering: I was just going to say, yeah, it passes. So you accept the Scientific Methods Panel

moderate rating for the composite construction. Okay.

Co-Chair Thraen: Terry, feasibility?

Member Fairbanks: Feasibility, this is just like I said. We have the advantage of new data that we're evaluating in the scientific support piece. We have the advantage of history of use here that demonstrates clearly feasibility.

This is being used all over the country, and its data elements are both electronic and data abstraction. It's noted it takes significant effort to do it. But it's clearly feasible.

Co-Chair Thraen: Discussant comments?

Member Roney: I completely agree with Dr. Fairbanks. And we did not pass it on feasibility when we voted there in Washington, D.C. But it's clearly feasible. It's just not all -- you can't get it all easily electronically. You do have to do some manual chart abstraction.

Co-Chair Thraen: Okay. Any comments or questions from the committee members?

(No response.)

Co-Chair Thraen: All right. Let's vote.

Mr. Sakyi: Voting is now open for Measure 0500 on feasibility. The option is A for high, B for moderate, C for low, and D, insufficient.

Co-Chair Thraen: That's Question No. 11. Because of accepting the Scientific Panel's recommendations, 10, I think, was skipped.

Mr. Sakyi: That is correct.

Mr. Sakyi: Waiting for one more.

Mr. Pickering: Is anybody having issues with the voting link? There we go. Thank you.

Mr. Sakyi: Voting is now closed for Measure 0500 on feasibility. We have 3 votes for high, 13 votes for moderate, 1 vote for low, and 1 -- and zero, insufficient. I apologize. The measure passes on feasibility.

Co-Chair Thraen: All right. So next -- is it up to use now, Matt?

Mr. Pickering: Yeah, yes.

Co-Chair Thraen: Yeah? Okay.

Mr. Pickering: Now we're on use.

Co-Chair Thraen: Use, Terry.

Member Fairbanks: This is a must pass criteria, I understand it. And it is used in accountability and payment programs. It's used by CMS and some states. There were no major issues with use that were identified at all.

Co-Chair Thraen: Okay. Discussants, any comments?

Member Roney: I would only like to give feedback to the developers on the usability. Because it is an all or none, once they hit a fail to meet the metric, some abstractors just stop abstracting in that record. And so it doesn't really lend well to quality improvement efforts to see where you fall out.

So, say you fall out on lactic acid within three hours, they don't keep abstracting after that. So you have no idea where your other areas of opportunity for improvement are. So one of the things I would like to see is it mandated that the charts are crafted so we can truly, truly lead our quality improvement

efforts at the institutional levels.

Co-Chair Thraen: All right. That was for the developer?

Member Roney: Yes, sir. Yes, ma'am.

Co-Chair Thraen: All right. Any other comments from the discussants?

(No response.)

Co-Chair Thraen: Any comments or observations from the committee members?

Mr. Pickering: I don't see any questions or hands raised. I'll just sort of add on that this is the use component. So again, use and usability, you see those together.

We're on use. And what we're looking at for use is if this measures use in some sort of public reporting or accountability application in which Terry had summarized as well. And we're also trying to see about feedback related to those being measured or others. And so feedback on this measure about how they could potentially use the measure to -- for some quality improvement as well.

The developer had provided summaries of how CMS collects information related to that. And so again, that's just use. And so we'll get into usability next. We just wanted to sort of remind folks that you see them together in these evaluations. But we're on use to see if it's used in a public reporting or accountability application which Terry had mentioned that it is, and then also looking about feedback from those being held accountable to measure or others.

Co-Chair Thraen: All right. Any questions?

(No response.)

Co-Chair Thraen: All right. We'll go ahead and vote.

Mr. Sakyi: Voting is now open for Measure 0500 on use.

Co-Chair Thraen: Question --

Mr. Sakyi: The options are A for pass --

Co-Chair Thraen: Sorry.

Mr. Sakyi: -- and B, no pass.

Co-Chair Thraen: Question No. 12.

Member Roney: Isn't it Question 13?

(Simultaneous speaking.)

Mr. Pickering: No, that's usability. We were --

Co-Chair Thraen: No, that's --

Mr. Pickering: -- still on use.

Co-Chair Thraen: -- 12.

Member Roney: Well, I answered the 13, I guess, team, sorry, already. If you can undo my answer, I don't know.

Co-Chair Thraen: So can you go back to 12?

Mr. Pickering: To the previous question?

Member Roney: I answered 12 and 13.

Co-Chair Thraen: All right. You're good then.

Mr. Sakyi: Okay. We have 17 votes. Voting is now closed for Measure 0500 on use. We have 17 votes for pass, zero no pass. The measure passes on use.

Co-Chair Thraen: Okay. And then just FYI, Matt, the survey is now holding on the last question that you -- it's working. You don't have to go through the front, start at the beginning and go all the way to the end each time --

(Simultaneous speaking.)

Mr. Pickering: Thank you.

Co-Chair Thraen: I used to talk -- whatever you fixed in the background, thank you. All right. Usability, Terry.

Member Fairbanks: Well, as a human factors engineer, I have trouble talking about this use of the word usability. But the preliminary rating is a high on this. The data demonstrates that the measure -- that the rates have improved over time and there are no major issues that were identified -- that I could identify or in the comments around usability. And I would just offer my co-discussants an opportunity if there's anything else I missed.

Co-Chair Thraen: All right. Any comments from the committee or questions?

Member Roney: I think it's just notable that it's a report only and not tied to financial consequences which to me is a good things as a clinician. I think it just truly can be used as quality measurement that way without fear that I also -- I think that the ability -- I don't think institutions would all focus on sepsis that much without this measure in place. And so it really does do the intent of trying outcomes through a national focus that I don't know necessarily -- some institutions would and some wouldn't. But I think this drives it for all.

Co-Chair Thraen: Thanks, Jamie. Any other comments?

(No response.)

Co-Chair Thraen: All right. Let's call for the vote.

Mr. Sakyi: At this time, we're open for Measure 0500 on usability. The options are A for high, B moderate, C low, and D insufficient.

Co-Chair Thraen: Question No. 13.

Mr. Sakyi: Voting is now closed for Measure 0500 on usability. We have 10 votes for high, 5 votes for moderate, 2 votes for low, and zero insufficient. The measure passes on usability.

Co-Chair Thraen: All right. And now I believe it's the final vote for endorsement.

Mr. Sakyi: That is correct. Voting is now open for Measure 0500 on the overall suitability for endorsement. The options are A for yes and B, no.

We have 17 votes in. Voting is now closed for Measure 0500, the overall suitability for endorsement. We have 14 votes for yes, 3 votes for no. With 14 votes for yes, the measure passed. The measure is therefore recommended for endorsement.

Co-Chair Thraen: All right. Thank you. Thank you to the developer, and thank you to Terry and his discussants for taking this particular one on.

Member Fairbanks: Iona?

Co-Chair Thraen: Yeah.

Member Fairbanks: Before the developers go, I just want to -- I think it's important to say that I feel like we're -- it's an important challenge that we heard right now because this was such a controversial measure in terms of the evidence. And

in the beginning, the developer said they were going to go back and look at revisions. They talked about the IDSA and recommendations, and they're going to look at antibiotic usage. They're going to look at fluid resuscitation.

And I think it's important to point out I became an expert in this over the last week with a deep -- and I think it's very clear that sepsis is one of our number one problems I have because of my work in patient and family advocacy. I have friends who have lost family members, children from sepsis and spouses. We cannot lose sight of this as a measure. So the developers need to come back to us next time having answered these questions because I hate to see this measure being in jeopardy.

Co-Chair Thraen: Thank you, Terry. I also lost my mother to sepsis. So I do think it's one that touches all of us in some way or another. And I do appreciate that this is an evolving measure. And I think that I won't be here, but I'm charging Don and John with the challenge of three years from now when it comes up for re-review that they take a look at the measure to make sure that the measure has incorporated some of these issues and is addressing the elements that have been discussed. So yeah.

Member Fairbanks: Thanks.

Co-Chair Thraen: Yeah.

Member James: May I ask a question, Iona?

Co-Chair Thraen: Yeah, yeah.

Member James: I think this is a fundamental issue. Can we tell the developers, here's a list of things you need to do and we want to see these when you come back to us and see that you have improved your measure?

(Simultaneous speaking.)

Co-Chair Thraen: Well, I think that's the role of the NQF staff as they collect and record our comments over the course of time. We certainly are giving them feedback in this discussion. The developers are always present -- usually present in this conversation.

And they're hearing the concerns as they've been expressed. And we'll take them -- reasonable developers will take them into consideration. And if they're looking to get past it three years from now, those would be the considerations that should be accounted for. I do think we've also asked NQF to record the concerns at each review, particularly when you're talking about maintenance reviews that the measures do evolve based on the feedback that they're getting.

Mr. Pickering: Yeah, John --

(Simultaneous speaking.)

Member James: This will come up again. But this is interesting. Thank you.

Mr. Pickering: No, I was just going to add, John, it's a great point. And there are instances where developers do take those recommendations that come from standing committee and other NQF convening groups to update their measures for reevaluation, whether it be for maintenance endorsement or coming back through for new endorsement if the measure does not pass, for example. NQF does provide us within our technical reports as far as those recommendations that the standing committee provides.

So the developers and the public can see what those recommendations are. And NQF also would incorporate any of those major challenges or

concerns that the committee has into subsequent measure evaluation meetings. And we've heard a lot today from those that have been involved previously with this -- specifically with this measure.

And some of the responses that the developer has noted may not have been all included into what we have today. But know, some of the issues with the measure as the standing committee has raised and also looking to make future updates potentially depending on what they find. In this case with 0500 is what the developer has mentioned. But John, you raise a really good point, and it's something that we constantly don't want to forget to do is recommendations made should be something that we document and carry forward for future maintenance evaluations.

Co-Chair Thraen: Thanks, Matt.

Member Roney: Matt, this is --

(Simultaneous speaking.)

Member Roney: -- Jamie. And I just want to give you kudos. I know that Dr. Rivers and Dr. Townsend remember the last time we did this. And we didn't have a scientific panel to look at the evidence ahead of time.

And it was a day and a half discussion where we got bogged down looking at the evidence and trying to come to consensus as a group and accusations flying that some of us didn't have the right skill set to even be evaluating evidence. I would say National Quality Forum changed a lot since our last endorsement of this measure. And I know that at least myself and our developers appreciate the fact that this was so seamless compared to last time.

And we really didn't have to dive into the evidence they brought to us. We just had to kind of look at

what the scientific panel said. And so kudos to NQF for making the process better, and I think that's just an example of how you saw the problems.

And it was probably our group that lead to that scientific panel honestly. Maybe not, but I'm going to guess so. It was a pretty heated discussion. So it was great work on your guys' part.

3621: Composite weighted average for 3 CT Exam Types

Mr. Pickering: Well, Jamie, thank you for the kind comments. I do want to make sure we continue forward with the 3621 measure discussion. But we really can't do any of this without you all.

And so it's your expertise that is really needed in this consensus forum. So thank you for the kind comments as well. I do want to move onto 3621 and I'll turn it back over to our co-chairs. Iona, I believe this is you taking this one, and --

Co-Chair Thraen: Yes.

Mr. Pickering: -- I'll see if the developer is on for this measure. Is the developer on the call, American College of Radiology?

Ms. Campos: Hi, Matt. Yes, we are.

Co-Chair Thraen: All right. Thank you. Just one quick note. I have a hard stop at five minutes before the hour. And so Ed, if I have to leave before we get through this, it'll be in your hands. Okay?

Co-Chair Septimus: I'll do my best. I am now unmuted.

Co-Chair Thraen: Thank you, Ed. All right.

(Simultaneous speaking.)

Co-Chair Thraen: That's right. All right. So this is a composite weighted average for three CT exam types. The steward is the American College of Radiology.

It's a new measure. It's a weighted average of three CT exam types, overall percentage of CT exams for which dose length product is at or below the size-specific diagnostic reference level for CT abdominal pelvis with contrast, single phase scan, chest without contract, single phase scan, and head/brain without contrast, single phase scan. With that, I'll turn it over to the developer to provide a summary of this measure.

Ms. Campos: Thank you very much, Iona. My name is Karen, and I'm representing the American College of Radiology this morning.

Co-Chair Septimus: Can you speak up? Karen, can you speak up a little bit? You're hard to hear.

Ms. Campos: I'm so sorry. Can you hear me better now?

Co-Chair Thraen: A little bit. You could probably scream and it would work better.

Ms. Campos: Okay. I'll do my best to speak as loudly --

Co-Chair Thraen: That's good.

Ms. Campos: -- as possible.

Co-Chair Thraen: You got it. You got it. Okay.

Ms. Campos: Thanks. So this is a composite weighted average for three CT exam types. This measure captures how well radiation exposure from the scanner is adjusted for patient size using size-specific exam level diagnostic reference levels and

how well total radiation exposure from an exam is optimized based on the CT dose index, dose length product, or DLP.

The measure goal is to decrease preventable harm through the effective optimization of CT protocols and reducing radiation exposure to patients. This measure was evaluated by the SMP in March. It passed with a moderate rating for reliability, but consensus was not reached for the validity portion.

Our original submission included two forms of face validity, consensus documents, and measure usage. The panel had requested a more formal version of face validity. So we conducted a face validity survey that was reviewed during the SMP meeting.

Some of the levels that the SMP had were with respect to the level of analysis that the survey analyzed. The survey was conducted for the composite measure and its components at both the group and facility level. Our risk stratification was only performed at the facility level because clinician groups typically support one or more facilities and calculating the measure for a clinician group is an aggregation of the measure as calculated for facilities.

Radiation dose optimization is conducted on each facility by the team on site. The overseeing clinical group may or may not apply uniform practices across all facilities they serve. Lastly, there was also some discussion on the last question of the face validity survey which received 62 percent agreement.

We didn't use a Likert scale for the survey which may have impacted the results since the participants were not able to indicate a degree of agreement. Several respondents who disagreed felt that DRLs were not designed for rating performance

which is true. DRLs were developed for quality assurance and identifying outlier cases and ensuring that system level opportunities for improvement such as protocol optimization are addressed.

Even so, they may be used as an alternative way of identifying performance gaps. We recognize that there are other important aspects in evaluating dose optimization. But this measure is one of the best ways currently available for evaluating facility performance for accountability. I would like to invite Dr. Mahesh, chief physicist and professor of radiology and radiological science at Johns Hopkins to speak more about DRLs and the applications in clinical practice. Dr. Mahesh, are you there?

Dr. Mahesh: Yes, thank you. Can everybody hear?

Ms. Campos: Yes, we can hear you.

Dr. Mahesh: Okay. If I'm too loud, please let me know because I'm on the opposite end of the spectrum with Karen. By the way, my name is Mahesh, I'm the professor of radiology and cardiology at Johns Hopkins University. And I'm also pivoting more in this field because recently I was the vice chair for the National Council of Radiation Protection which came out with a report called radiation exposure to the U.S. population.

What I'm going to talk about is three things. I'm not going to take too much time. First of all, the CT use in the U.S. has increased and going up which is a good thing because we are doing a lot better in imaging and diagnosing. Just to give an example, in 2000, we were doing approximately 35 million CT procedures in the country.

By 2019, that number has gone up to 91 million procedures. In 2020, the number went down but because of the pandemic. Anyway, the reason why

dose optimization is important is the technology is helping to a large extent going down in the dose.

However, the usage makes a difference because how the user is setting up the protocol. And this particular measure is one way to help the facility to see where they stand. So why did we chose the diagnostic reference level?

Diagnostic reference level, as Ms. Karen said, it is typically used for quality control. However, in a clinical sense, that's the best opportunity we have to see how a facility does -- is doing with respect to the diagnostic reference level set on a larger scale. So it's not -- so it provides an opportunity for the facilities to check how they are doing in the representative set of cases.

The reason why the chosen three of these, like the CT of the head, CT of the chest, and CT of the abdomen pelvic, because they constitute the largest portion of the CT scans done in the country. As I mentioned earlier, as we were working on the national report, we had looked at all the procedures and head, chest, abdomen, and pelvic CTs account for nearly 75 percent of all the CT procedures done in the country. Therefore, if this as a representative of CTs used to evaluate the site, if the site look at these DRLs and compare to the standard DRLs, that automatically helps to look at all the protocol, minimize the radiation as much as possible, still obtain good image quality.

That is the idea of the whole thing. And just wanted to tell you this measure has already been put in use. It's already been used by the CMS. All we're requesting is I think this is a fairly good measure to evaluate a site. It's not the ultimate best one. But in the absence of any other measure, this is a simple way for the site to compare itself how they are doing. That way, they can actually reduce it. The

best analogy I tell in my classes, like, you have different variety of car, and it all depends on how you drive your car, from the slowest to the BMW.

The driving method is what the protocols I'm looking here. And if the site knows what the protocol is doing, then they can actually adjust the protocol to come below the DRL. I think that's the reason why we considered this quality measure is a good measure, not too onerous on a facility but easy to capture and easy to provide. And they can demonstrate that by monitoring this, they can actually help doing a better scan for their patient. I think I'll stop here. And I will be happy to answer any question.

Co-Chair Thraen: Thank you for that. All right. I'm going to turn this over to Don. He is the lead discussant with Robert Green, Laura Kinney, and Nancy Schoenborn with a recusal from David Seidenwurm. Don?

Member Yealy: Thanks very much. It's nice to finally engage after an hour and half of silence. So as was noted already by Iona as well as the measure steward, this is a new measure and is a weighted average of three very common CT exams.

I have to tell the steward when I read that this was -- I forget what you said, 75-80 percent. In my experience, I would've guessed a higher number than that. The number of things that this excludes is actually very low.

And it looks at the use of varying dosing imaging in these very common exams under the premise that higher is not always better. And by better, I mean it does not necessarily produce a better image. And we know that while for a singular image for there to be either lack of benefit or harm, the dosing would have to be dramatically high.

But we know that over time, exposure to ionizing radiation has both theoretic and then some real time concerns. And we'll get into that conversation as, what is really the patient-centered outcome on this even though it's a process measure? So you've gotten a fairly good description, and it's a fairly straightforward and simply measure to understand that we're going to be evaluating. I think, Iona, what we start with first is the evidence criteria. Is that what you'd like me to get into?

Co-Chair Thraen: Yeah, correct.

Member Yealy: Okay. So this has been through the preliminary ratings and notes. And what we know is that the evidence as presented by the steward is that there's variation in the diagnostic reference levels that are used for these three different types of images. And this included direct evidence and systematic review.

This is actually a process outcome. As you've heard on the evidence side, most of the conversation about a concern will center on while there's variation, is a diagnostic reference level the appropriate tool to use and does that tie to an important clinical outcome? It's okay to have a process measure. But we, in general, have to believe that it actually ties to something that matters as opposed to it does not.

I do not think that, and I will invite the steward -- I do not think we'd be having this conversation about people with isolated imaging because I don't think there's a lot of strong evidence behind that. But we know that there's sub-populations in whom imaging will happen repeatedly over time.

And secondly, those who have the opportunity for ionizing radiation, not only from this but from other things, in other words, no one would really have a

lot of thought about a 75-year-old undergoing a singular isolated first-time scan unless there was some horrific kind of event. So that's where we're at with the evidence.

And the preliminary rating was that despite those concerns, there was a moderate level of evidence to support the use of it. And there was a recommendation for passage at that level. I'd invite now, if I can, Iona, the steward to comment on the DRL as well as the length of the process from an evidence perspective to an important patient-centered outcome.

Co-Chair Thraen: Absolutely.

Ms. Campos: Thank you. I would like to clarify. This is an intermediate outcome measure, not a process measure. So each component of this measure captures how well radiation exposure from the scanner is adjusted for patient size using size-specific exam level diagnostic reference level and how well the total radiation exposure from an exam is optimized based on the CT dose index DLP. So it's considered an outcome because it's an assessment of the effect of the healthcare system on the patient. And if you have any questions on the evidence -- specific questions on the evidence that we could answer.

Member Yealy: Yeah. So it's actually an intermediate process, not an intermediate measure. So it's not an outcome measure, per se. I think we're saying the same things. So the question is, again, if you want, if the preliminary rating was moderate. If you have any thoughts about the DRLs being appropriate, that's already been -- you've touched upon it -- I don't know if you have anything else you want to add on it -- and the link to a very specific patient outcome or series of outcomes.

Ms. Campos: Dr. Mahesh, would you like to speak to that?

Dr. Mahesh: Sure. I'd be happy to. Thanks, Don. And you kind of touched on the point, important point especially, that a few -- like, frequent flyers. A few patients receive a lot of CTs even though a single CT scan of a person is not a big deal. But a few group of the populations have had a number of CTs. And first, it is not adjusting the protocol to their size and to their technique we're unnecessarily exposing the patient for no benefit for the patient. In a way, we might be even be doing harm if we do a lot of these at high doses.

And in that regard, the DRL's widest commodity is because not only here in the U.S., both in Europe across, they're introducing this DRL as a good reference point for a site to evaluate how they are doing. In that regard, if we use that, the site will be able to adjust the protocol and deliver a better dose, whether it's a single dose to a patient or 100 scan to one patient. Either way, it could impact the whole thing. That's the idea.

And I agree with you, 75 percent is what we found in a national survey of the most commonly done, chest, abdomen and pelvic CTs. If it is only a -- like 100 percent -- like, 90 percent is chest, pelvic and abdomen, so that number is pretty much we know the urgency of this particular measure.

(Simultaneous speaking.)

Member Yealy: Was there any consideration ever -- and I realize this is about targeting those in which the image was something greater than the first or the first in a time period because, I mean, I think we agree complete. In other words --

(Simultaneous speaking.)

Member Yealy: -- a much more targeted surveillance --

(Simultaneous speaking.)

Member Yealy: -- because I think that would eliminate a lot of the concern. And I wonder if as you were developing it if that was a thought. I realize the burden that does --

Dr. Mahesh: No, I'm not sure about that as a point here because that is an important question the whole field has to answer. Like, Hopkins we sometime -- and we had the same patient come back, we know whether we use the right protocol to make sure that the doses are not high. I think Karen can answer, but I don't think that is the emphasis of this particular one because that will take a whole ballpark of a study to follow a patient and then see whether the techniques were different and so forth.

But I was enrolled in another study a few years ago where we looked -- randomly looked at the dose data across four to five institution. Even within and outside institution, we saw almost three to ten times dose variation. And that is a disturbing factor because even though a head CT is a head CT, different facility will do different way. But having a measure like this will help them to close the wide margin, three to ten, to much more closer. That is one of the main benefit of this particular measure that a facility can benefit from.

Member Yealy: Okay.

Member Sood: This is --

Member James: May I ask a question, please? This is John James. So I'm trying to understand this. So the idea of limiting the exposures is to prevent cancer downstream primarily. Is that right?

Dr. Mahesh: That's -- that's -- the idea -- that's correct. The idea is, like, we want -- there is no point in exposing a patient more than necessary because theoretically we don't have much evidence at the one or two scans. But if you do hundreds of scans, there is some evidence from the survivor of Hiroshima and Nagasaki population that can increase some risk for the cancer.

However, rather than worrying -- in order to make sure, we have a model application to make sure all the scans are done properly, not too high. That is the main emphasis because we want to make sure we use it as an ALARA principle, as low as reasonably achievable principle, to get good images with the right kind of dose. And facilities -- majority of the facilities are doing it, but they don't know how they are doing.

But if they compare this type of a measure measurement and then they will see where they are. If they are already there, we don't have -- they don't have to do much further. But otherwise, they can fine tune the protocol to get the right dose. That is our main goal of this particular --

Member James: Okay. My question is, so institutions are supposed to have their own protocol. But there are so many uncertainties here, what cancer risk is tolerable and how old the patient is and how large the patient is and how many scans --

Dr. Mahesh: Yes, so --

Member James: -- should be done. I'm afraid there's a lot of uncertainty here and quality.

Dr. Mahesh: Understand.

(Simultaneous speaking.)

Member James: -- with higher doses. Is this kind of an expert consensus? Is that where the gold standard protocols come from?

Dr. Mahesh: There is no gold standard protocol. There is always variability depending on the radiologist, perceptibility of the image quality. But generally, there's a ballpark protocol which can be used.

To answer your question about the cancer risk, that's another discussion for almost I can speak for one hour with all the controversy, plus and minus, because at the low end, we don't have good evidence about the causing cancer. But with repeated number of scans, there is a good possibility that can trigger cancer. And then again, this particular measure also take into account the patient size also for the first time which would tell you that a large patient does not -- may need a higher dose to penetrate the body to get a good image quality.

The same technique should not be used on a thin patient. And that will kind of correct in these things. There's a lot of uncertainty in the dose measurement which we physicist know. However, it is in the absence of any other measure, this is a good starting measure because the dose metric put out is from the machine. The machine output is what the CT volume and then the length. So that will kind of very much kind of easily tells us how a facility is doing these procedures.

Member Sood: This is Geeta. Thank you for that description. And I think Dr. Yealy answered a big part of my question. I understood his suggestion to be limiting the denominator of who we're assessing radiation exposure on which certainly seems to make sense and, like you said, was more difficult to do. I would imagine and I believe that we know that

less radiation is better, generally speaking. My question is around what you said related to the uncertainty of what -- how much radiation is needed. So you had mentioned that body size determines how much radiation is needed.

And my understanding of what the Scientific Methods Panel had some concern about is, are we comfortable that the amount of radiation that this metric thinks is needed is, in fact, reliable -- is, in fact, what would be needed for high resolution CTs? Does that make sense? I'm not sure if I --

Dr. Mahesh: No, it absolutely makes sense because the reason is, like, we don't want to go too low. We don't want people to go too low because then images look so grainy in the patient. The radiologist will not be able to diagnose. At the same time, we don't want to go to high where the higher -- high -- very high quality does not make any difference. So we want to use it the right level which has been established because of all the studies being done. All the radiologists are using the images.

And that reasonable number is what the DRL tells. So if a facility is using higher than that, this gives an opportunity to pull it down. We are -- I'm very much against going too low because that jeopardize the patient. So we don't want to have that as a trial. We want this to be come down at the diagnostic reference level. We are not setting a base level, but the right level is what we are talking about because we want -- I appreciate because the image quality is the most important one.

And if I want to get my kid get a CT scan done, I'm more worried about the right diagnosis than the radiation dose. And again, but I will make sure that the facility using the right protocol because I don't need to expose my kid for a higher dose. They can just -- they need to use the right dose. So a lot of

the time, I tell the patient to ask the facility what protocols they're using. By just ask, we have found, many times, they are more careful in doing the CT. And these measures will help us.

Co-Chair Thraen: I'm going to stop you there and ask if the discussants have any other comments they'd like to make.

Member Yealy: I do not have any more, Iona, as the lead discussant. Again, the question is, is the application of this across all who receive imaging the ideal way to address the problem or should those who have a different background threshold for having the side effect we're talking about? And that's going to be those with repeated measurements and/or who are young is proportionately where it can happen. But that's the concern I have regarding evidence. And we've already had the conversation about that.

Co-Chair Thraen: So either Robert, Laura, or Nancy, do you have any comments?

Dr. Mahesh: Can I make response to that?

Co-Chair Thraen: No, I'm going to ask you to hold.

Dr. Mahesh: Okay, sure.

Co-Chair Thraen: Thank you. Any other questions from the committee members before we take a vote?

Member Falvey: Iona, this is Jason. I had a quick question. I'm sorry --

(Simultaneous speaking.)

Co-Chair Thraen: Go for it.

Member Falvey: So I had a question along the

similar lines. I'm thinking about trauma exclusions and if you guys had considered that at all because I think there could be a potential risk benefit issue here. And I don't know if that's in your data set or something you had access or considered from a top level in terms of an acute stroke or something that might require some urgency and may not be able to kind of optimize that radiation dose in a time efficient manner.

Dr. Mahesh: May I respond?

Co-Chair Thraen: Yes, that's directed to you.

Dr. Mahesh: Okay. That's a good question because that is also included in the (audio interference) that can offshoot. But typically, if an institution provides all the data, look at all their data, the number of trauma patient studies will be much lesser which is not going to impact. But to answer your question, as an academician, I do agree that including that in this measure can offset a site and because the trauma is, like, it should be excluded on, like, profusion scan.

We're only looking at the CT of a head. That's why we're only limiting -- this measure limits to only three of them as CT of the head, chest, and abdomen with contrast. There are other more protocols which we are not looking, so that can be a trauma protocol or a dual-energy protocol or a CT profusion protocol or a cardiac CT profusion -- we are not looking to them. These are the more common one. As I mentioned, we hold, like, nearly 75 percent of all the CT use CT procedures done in the country is these three category. That's one of the reason why this is being used.

Co-Chair Thraen: Does that answer your question, Jason?

Member Falvey: Thank you.

Co-Chair Thraen: All right. Any other questions?

Mr. Pickering: So Iona, I'll note maybe Don, if there's any public comments related to this measure within -- that have been submitted? Maybe we could touch on those, similar of how we did with the other composite, 0500. And I also will mention again that this has been submitted as an intermediate outcome measure.

So according to NQF processes and how we evaluate this for evidence, it's considered to be similar as a process measure when it's an intermediate outcome. So you're looking again to see if those processes -- or what has been listed within this measure -- the processes lead to some sort of desired health outcome. And this is a composite, so each one of the components should have supporting evidence as well.

So similar to how this was for 0500 where that was explicitly process components, this one is deemed as an intermediate outcome. But as an intermediate outcome, according to NQF criteria, it is looked at for evidence and treated similar as a process, looking for a connection to a desired health outcome. So maybe, Don, I don't know if we wanted to look at the comment that was received --

(Simultaneous speaking.)

Member Yealy: Yes, so the comment received, I think we've touched upon, it's just -- it's worded much more eloquently and from the view of someone in the field. It really focuses on the fact that really no one disputes that unneeded exposure to ionizing radiation doesn't serve a benefit and has real potential harm. The size of that is something that we could have a long conversation and which

people are likely to do it.

But whether or not this measure as constructed identifies the most effective process that could link to that particular outcome and modifies the behaviors that are most important which is facility protocols, individual radiologist choices for specific imagining sequences, I really do think these are different views on the same comments that we've actually been exchanging back and forth and get us back to the right series of summative measures to produce the outcome link that we're hoping for.

Dr. Mahesh: Can I comment?

Member Yealy: Actually, no. I was not going to ask about that, unless -- because I think we've discussed this. I'm just sharing that that goes along with what I think we've had conversation about.

Dr. Mahesh: But that particular specific comment, I would like --

Mr. Pickering: So -- sorry for the developer. Yeah, unless the standing committee is requesting a clarification, we're just going to reserve any further comments, unless the standing committee would like any clarification from the developer.

Co-Chair Thraen: Thank you for bailing me out, Matt. How about we call for the question? And after this question, I'm turning it over to Ed. It's his turn because I have to step off.

Co-Chair Septimus: I'm so excited, number one, to be able to speak. And number two, I can now vote.

Co-Chair Thraen: Yes. So should we go ahead and start the survey?

Mr. Pickering: So does anyone else from the standing committee have any other questions or

comments?

Co-Chair Septimus: Oh, it went back to the beginning again.

Co-Chair Thraen: Yeah.

(Simultaneous speaking.)

Co-Chair Thraen: It's also a new measure.

Mr. Sakyi: Yes. So when you refresh the browser, please make sure you are seeing 3621 composite weighted average --

(Simultaneous speaking.)

Co-Chair Septimus: What --

Mr. Sakyi: -- next question.

Co-Chair Septimus: What do you mean refresh my browser? I mean --

Member Sood: It doesn't actually have the name of the measure, as least from what I'm reading. It just says, Question 1 of 13.

Co-Chair Thraen: If you refresh it, Geeta, it goes back to the first page which says 3621 composite weighted average.

Member Sood: Got it.

Co-Chair Thraen: It should. And then you start the survey. Refresh is that little circle arrow on the top left-hand side, Ed.

Co-Chair Septimus: Okay. Well, I'm going to have to sign on again. This is -- as I said, if this is an enhancement --

Co-Chair Thraen: No, it's in your URL.

Co-Chair Septimus: No, I got it. I got it. I just have to --

Co-Chair Thraen: Okay.

Co-Chair Septimus: -- re-sign on. That's all. That's easy. That's not a problem.

Co-Chair Thraen: Okay. All right.

Co-Chair Septimus: So it's just --

Co-Chair Thraen: So we're on Question 1 of this new measure.

Mr. Sakyi: Voting is now open for Measure 3621 on evidence. The options are A for high, B for moderate, C for low, and D, insufficient. We are expecting 19 votes.

Mr. Pickering: Right. So again, that's David Seidenwurm is recused from voting. So out of 20 that we originally have, we should be expecting 19. And then Iona will probably be dropping off which we would expect 18 moving forward, so --

(Simultaneous speaking.)

Co-Chair Thraen: Thank you, guys. See you tomorrow.

Mr. Pickering: Thank you.

Co-Chair Septimus: Bye-bye. Hopefully not tomorrow, Iona. We'll see.

Co-Chair Thraen: Okay. That'd be good. All right. Take care. Bye.

Mr. Pickering: Okay. We're still missing one more. Ed, were you able to get back in?

Co-Chair Septimus: Yeah, I voted.

Mr. Pickering: Has anyone else not voted or having issues? Iona said that she voted. Okay.

Member Yealy: Got everyone, Matt?

Mr. Pickering: Not everyone. It looks like we're still missing one -- there we go. Okay.

(Simultaneous speaking.)

Member Yealy: You're not going to ask them to raise their hand, are you?

Mr. Pickering: Go ahead, Isaac.

Mr. Sakyi: Voting is now closed for Measure 3621 on evidence. We have zero for high, 15 votes for moderate, 3 votes for low, and 1, insufficient.

Co-Chair Septimus: Well, that passes on.

Mr. Sakyi: The measure passes on evidence.

Co-Chair Septimus: On evidence, right?

(Simultaneous speaking.)

Mr. Pickering: Yes.

Co-Chair Septimus: Okay. Don, do you want to go on to the next section?

Member Yealy: Yeah, next is to evaluate the gap. This received a preliminary rating of moderate which is essentially expected for this type of measure. And we were supplied with performance rates for the measure that spanned 2017 to 2020 and identified that there are opportunities. Those have -- those seemingly align with not only my personal clinical experiences in an oversight role but the evidence that exists from the views of specialty societies and other potential experts. There weren't disparity data in this. I don't think that that, by

itself, is necessarily an issue.

I think dosing is an equal opportunity offender and not likely something directly affected by that and likely picked up in other parts of the ordering and deployment processes. So there were no significant issues with the gap analysis on the preliminary part. And I did not see comments focused from the external commentators or inside the group that would actually alter this particular assessment now.

Co-Chair Septimus: Laura, Nancy, any other comments from your perspective? And then we'll ask the committee.

(No response.)

Co-Chair Septimus: Any comments from the committee?

(No response.)

Co-Chair Septimus: Then I say we go ahead and we vote on performance gap. And as that's being put up, we're getting close to what would've been the lunch hour. I suggest that we finish this measure and then we break, unless somebody is absolutely so hungry that they can't wait.

Mr. Pickering: Sounds like a good idea.

Co-Chair Septimus: Yeah, I'm going to have to sign back in. I got kicked out again. Oh, gosh. Okay. Let's try again.

Mr. Sakyi: Voting is now open for Measure 3621 on performance gap. The options are A for high, B moderate, C low, and D, insufficient. And we're expecting 18 votes now.

Okay. We have 18 votes. Voting is now closed for Measure 3621 on performance gap. We have zero

votes for high, 18 votes for moderate, zero for low, and zero, insufficient. Measure passes on performance gap.

Co-Chair Septimus: Thank you. I think this is a composite measure, right, Matt?

Mr. Pickering: That's correct. That's --

(Simultaneous speaking.)

Co-Chair Septimus: All right. So the next thing we have to do is vote on rationale. Don, anything to add on that beyond what we've already discussed?

Member Yealy: I'm sorry. Are we doing reliability now or --

Co-Chair Septimus: No, no. It's a composite measure, so we vote on rationale.

Member Yealy: I don't think there's any new information. And we've discussed -- this is one of the facets that often overlaps with other parts of the conversation. Much as we try to cleave them, they end up getting married all the time anyway. And I think we addressed that issue already about how does this tie to the specific things that were perhaps the most interesting.

Co-Chair Septimus: Laura, Don -- I mean, Laura and Nancy?

(No response.)

Co-Chair Septimus: Okay. Committee members, any other discussion?

(No response.)

Co-Chair Septimus: Seeing none, I suggest that we go ahead and vote, Isaac.

Mr. Sakyi: Voting is now open for Measure 3621 on the composite quality construct. The options are A for high, B for moderate, C for low, and D, insufficient.

Mr. Pickering: Similarly with the previous measure, this is really thinking about whether the composite construct has good rationale and is articulated well and is logical. And so that includes how it is calculated. So just a refresher on that.

Mr. Sakyi: We're waiting for one more vote.

Mr. Pickering: Ed, did you get kicked out in this question too?

Co-Chair Septimus: I voted, then I got --

Mr. Pickering: Okay, good.

Co-Chair Septimus: -- kicked out again. But I voted. Why did I get kicked out again?

Mr. Pickering: I'm not sure what's going on, on your end. I'm sorry, Ed.

Co-Chair Septimus: No, but I voted. I voted. But I just don't understand why I got kicked out again.

Mr. Sakyi: Yes, we have 18 votes. Voting is now closed for Measure 3621 on the composite construct. We have 2 votes of high, 14 votes for moderate, 1 vote for low, and 1, insufficient. The measure passes on the composite construct.

Co-Chair Septimus: Thank you. And so Don, we'll go to scientific acceptability of the measure properties, first going on reliability.

Member Yealy: I apologize. And I actually got a little ahead of this and I apologize. This is another must pass criteria. Again, as is common for these types of

measures, not only did it go through preliminary review but scientific review. The preliminary rating on reliability was moderate, and that's the upper end of a receivable score for this.

The reliability score for the measure as it sits now which is a different conversation than what we just had and what we're about to have is actually exceptionally high. And there was no dialogue about this, either amongst the panel, the NQF panel or for those who are inputting externally. And so I don't think there's a lot to say here unless there's other questions from the discussants or from the NQF committee.

Co-Chair Septimus: Laura and Nancy and anybody else on the committee?

Mr. Pickering: Looks like Jason has his hand raised, Ed.

Co-Chair Septimus: I'm sorry, Jason. For some reason, I can't -- I don't see you. Okay. Jason, go for it. Thank you, Matt.

Member Falvey: No worries. There was one concern raised in the measurement panel. And me as a statistical person that does statistical type of research, the exceptional high rating for reliability did raise some validity concerns. So I know it kind of melds both worlds here.

But I think my concern might stem from this is drawn from a voluntary database from the American College of Radiology which might be a more homogenous sample than a nationally representative sample of hospitals. So if you have high reliability between invested and motivated hospitals who are voluntarily reporting this data, I don't necessarily know if that is what we're going to see on a national basis. So I think that's a concern

we should at least have in the back of our mind and certainly something we're going to ask when and if this developed measure comes back to us in the future.

Member Yealy: And Jason, I had the exact same thought and really did not focus much on it because I'm almost uncertain how you could -- at this stage of a new measure how you could truly address that. But I think you're right. In four years, I think we're probably sending a shot across the bow to the steward assuming that it gets through that this will be important to know. And like you, I would expect the reliability score to change. The real question will be not does it change but does it remain above a threshold that allows us to continue?

Co-Chair Septimus: Good, good point. Any other comments?

(No response.)

Co-Chair Septimus: Okay. Then I guess we're ready to vote on reliability. This is another must pass.

Mr. Sakyi: And we're voting on accepting the Scientific Methods Panel's rating.

Mr. Pickering: Right. So as I said, this is SMP --

Co-Chair Septimus: Correct.

Mr. Pickering: Right. The standing committee has been asked to accept the rating.

Mr. Sakyi: Voting is now open for Measure 3621. The question is, do you accept the Scientific Methods Panel's high rating for reliability? The options are A for yes and B for now.

Voting is now closed for Measure 3621. We have 94 percent voting yes to accept the Scientific Methods

Panel's rating and 6 percent no. The standing committee votes to accept the Scientific Methods Panel's rating for reliability.

Co-Chair Septimus: Okay. So then I guess we go to validity. And this is another must pass. And since this is a new measure, and Matt, correctly if I'm wrong, this is going to be based on face validity since it's a new measure and there's no, obviously, experience with the measure. Is that correct?

Mr. Pickering: So you are correct, Ed, that face validity is an acceptable form of validity testing for new measures.

Co-Chair Septimus: Right.

Mr. Pickering: Right. So the -- and the developer had provided that face validity testing. However, I'll just remind the group that the SMP again evaluated the measure. They had consensus not reached on validity. So the standing committee will be asked to provide their own rating as opposed to upholding the Scientific Methods Panel rating since they have consensus not reached.

Co-Chair Septimus: Right. That was -- thank you. That was going to be my next comment, that they could not reach consensus on that. Thank you. Don?

Member Yealy: Yeah, and to go into this more, I think this is almost a nuanced layered approach to the validity must pass criterion. And while the Scientific Methods Panel had a majority approval, it did not reach consensus as the existing rules. And if you look at this in its layered thing, virtually no one disagreed about the importance of monitoring radiation doses.

This reminds me of the conversation that you just had -- I can't say that I had -- about sepsis, although I've had it a bunch of times outside your

presence. That is sepsis is an important thing. So we can understand that radiation doses are important and useless added radiation doesn't serve anybody particularly well.

Majority of the panel, 71 percent, agreed that the components as described were reasonable and were one appropriate way to assess performance quality. Where things begin to degrade is that a very much smaller majority, at that point, 62 percent agreed that the scores would differentiate clinical performance across providers. And members of the Scientific Methods Panel also questioned the level of analysis.

The clinician group versus the facility, again, this calls into play exactly who influences the performances, not so much even on an individual scans but on collections in groups. And where does that happen at? And is the analysis done at the appropriate levels? According to one of the SMP members, the measures use the current measure as the CMS does. It aligns well with that. But we usually look for a more formal process than that.

The developer has also noted that they had lots of consensus documents. I don't think there's a lot of disagreement about that. The use of the measure outside of our particular evaluation of this has increased over a year or so. There must be some underlying strong face validity because people are adopting it and that the risk stratification is performed at the level of the facility and not at the group because the developer stated that the groups are generally aggregations of the facilities.

And a group often, if not usually, supports more than one facility in this day and age and that any findings on patient size stratification applicable at the facility level would be applicable at the group level. I guess I would ask the developer, have I

summarized your responses to this particular concern well? And if not, if you can be pithy about the reply where I may not have -- where my language may not have been precise enough.

Co-Chair Septimus: Karen or -- you want to take that, Karen?

Ms. Campos: Thank you so much, Dr. Yealy. That was truly excellent. I do want to just add a couple of nuances to that. For the face validity, I think I mentioned in my intro, we didn't use a Likert scale. So it was a yes or no, and respondents couldn't adequately tell us how much they agreed or disagreed. And I think that did influence our responses on the last question.

And I did want to point out that 100 percent of the patient and non-clinical survey respondents did agree that this was a good, worthwhile activity to determine from good or poor quality performance. So that was 100 percent there. And then in regards to the nuances for whether we should use diagnostic reference levels for rating performance, we still feel that they can be used to identify performance gaps, and it's one of the best ways that we have right now to do that. And Dr. Mahesh can very more eloquently explain that than I do.

And again, with the CMS approval, I mean, it is also a rigorous process to get that measure in an accountability program. And those stewards did feel that they were -- this was a good, worthwhile measure to use that for. So thank you.

Co-Chair Septimus: I think you did very well, Karen. Dr. Mahesh, do you have any comments on what Karen just said?

Dr. Mahesh: I think as we can -- less is better. So I don't want to cover anything else. I think she said

very nicely.

Co-Chair Septimus: Well said. Okay.

Member Yealy: Wow, I asked for 50, but that's really something.

Co-Chair Septimus: All right. John, you have a comment.

Member James: Thank you. Thank you, Ed. Well, I'd like clarification, okay. So I assume this excludes pregnant women. And does it take into account multiple CT scans, because those that get -- well, obviously more CT scans should perhaps have more cautious use of radiation in these things. How are those taken into consideration?

Dr. Mahesh: So Dr. John, I think I -- can I comment, Ed? This is a wholesome measure, like, taking everything coming into the facility and seeing how they are doing. I understand the concern which you have about that patient which are a lot of patient scans and pregnant patient.

That's a different one because this is just a first step in the way because that request a lot more and that can be a next measure to fine tune more closely into the practice. But this is giving a general what the practice is doing across the board. Let's say we're doing one million procedures -- safety procedure at Hopkins here.

The number of pregnant patient we scan is about less than 100, let's say. So there comes in the denominator comes as a washout. So as a first measure, we're seeing, like, across the board how we are doing for the head, chest, abdomen, and pelvic. So later, there can be other measures can be done, looking only at the pregnant patient or looking at only the prepared scans. So this will give you a general measure to tell how the facility is

behaving.

Co-Chair Septimus: So I think validity, we want to know the results, results about the quality of care will be better when implemented.

Dr. Mahesh: Yes.

Co-Chair Septimus: That's kind of the question we're trying to -- we're trying to answer in terms of voting on validity.

Member James: This seems to be like some homework assignment, I think, for the developer. And I think the developer plans that. So as long as the NQF lists these and says okay, here's things you've got to look at, I'm okay.

Dr. Mahesh: Thank you.

Member James: So basically, it's a process, an evolutionary process, and I'm having a little trouble accepting that, but that's reality. So thank you.

Co-Chair Septimus: Other comments?

Member Falvey: Ed, it's Jason. I just have one quick clarification. The other piece of validity is also the exclusion, like are we including and excluding the right people? So is including every single person that gets a CT scan, is that the right approach to start?

And I think that's a reasonable question. And I think trauma centers are going to have a very different population than rural community hospitals in terms of make up of what scans are being used, even if in the aggregate, the national average is different. So I think there has to be some careful consideration of exclusions and I'm not sure including every single person, and I still have my trauma concerns, working right across the street from a trauma

center, that I'm not necessarily sure that that is an appropriate use of an emergency department physician's time to quickly judge a patient's size when they think they're having an acute hemorrhagic stroke.

Member Yealy: So Jason, it's Don. As the emergency physician in the group now I understand what you're saying. This came up in the chat box.

It's natural in these conversations we try to bucket things into the headings that NQF walks us through. And no matter how much we try to bucket them, we end up talking about the issue just from -- maybe blind men around the elephant, and we see the same issue when we touch different parts of the elephant and -- in different ways.

I understand exactly the concern here now and I think it looks very much like the evidence conversation that we had before for very similar reasons. There's just different triggers to it. And it gets back to is this big lump the appropriate thing or much more focused. And you just have addressed it from clinical conditions versus you could do it from age.

One other comment, I sure hope no one is using the emergency physician's estimated body mass. I can promise you that's drivel and made up. We're not within ten stones of the right number, just to be a little snarky about it.

Participant: I'll endorse that.

Member Yealy: And I invite the developer, again, if you have something very specific to it. I think we've heard kind of the views on it before.

Dr. Mahesh: No, I think that you say it correct because you're looking at overall, there is so much variability in size. So I agree with you.

Co-Chair Septimus: Okay, are we ready to then vote? Okay, seeing no hands, I say let's go vote, Isaac.

Mr. Sakyi: Voting is now open for measure 3621 on validity. The options are A for moderate, B for low, and C, insufficient.

Mr. Pickering: And just to add to why only moderate is because this only has face validity. So to add any sort of empirical validity testing at the measure score level, we would see high. But since this measure is coming in with just face validity which again for new measures is an acceptable form of validity testing, this is why the options are limited to just moderate and low and insufficient.

Member Sood: Matt, this Geeta. Can I just ask a clarification question about how and what exactly we're answering for validity in this kind of --

Co-Chair Septimus: Geeta, we need to finish the voting and then you can follow. Let people finish the vote, then we'll come back.

Member Sood: Well, it actually has something to do with this specific vote, Ed.

Co-Chair Septimus: Well, Geeta, I hate to sound -- we've got to keep the process. We've already closed this to vote.

Member Charbonneau: What number question are we voting on, because mine has high. Is this question 5?

Member Yealy: No, it's question number 6.

Member Roney: And I got kicked out so I'm trying to get back in. I apologize.

Mr. Pickering: And Geeta, is this a content validity

question or more of a question on the criteria related to validity?

Member Sood: It's just a follow up on the clarification that you were offering about what -- when we're -- what exactly is this question? We're asking is the -- does the measure validly approximate the differences in radiation? Is that -- am I understanding that correctly?

Could you just put that into words what we mean by validity for this particular metric?

Mr. Pickering: Right, so in this case since face validity has been conducted, what you should be thinking about here is can this measure discern between good and poor quality.

Member Sood: That's beautiful and that's exactly what I needed. Thank you very much.

Co-Chair Septimus: That's what I said.

(Laughter.)

Co-Chair Septimus: Okay, do you want to vote? I hate to do this again.

Isaac, should we just vote all over again and make sure everyone is on the same page, is that okay?

Mr. Sakyi: Yes, we can do that.

Co-Chair Septimus: Let's vote again, because I don't want -- okay, so let's clear the responses. Tell us when we can vote again.

Mr. Sakyi: Voting is now open for measure 3621 on validity. The options are A for moderate, B for low, and C insufficient.

Co-Chair Septimus: That's question number 6, correct?

Mr. Sakyi: If you can still see your previous vote, you will need to refresh your browser.

Mr. Pickering: And Isaac, can you say the question number?

Co-Chair Septimus: It's 6, isn't it? Whoops.

Mr. Pickering: He has to go out of it.

Co-Chair Septimus: Oh, okay. He has to do it again.

Mr. Sakyi: It's question number 6.

Co-Chair Septimus: There we go. It's 6. Thank you, Isaac.

Mr. Pickering: So again, we're looking for 18 votes and we have 15. Anyone not voted? There's 16.

Member Roney: Sorry, one of them was me. I was still trying to get in again since I got kicked out, it took that long. So, great discussion.

Mr. Pickering: No worries.

Member Roney: It was very helpful to give me time.

(Laughter.)

Mr. Pickering: Thanks, Jamie. So we have all 18.

Mr. Sakyi: We have exactly 18. Voting is now closed for measure 3621 on validity. We have 12 votes for moderate, 3 votes for low, and 2 for insufficient. With 12 votes of moderate, the measure passes on validity.

Co-Chair Septimus: Okay. So I'll turn to Matt, since this is a composite measure, I think we have to vote on the quality construct again.

Mr. Pickering: That's correct. It wouldn't be voting

necessarily on it again, it's just voting on the empirical analyses or any of the analyses related to the composite.

Co-Chair Septimus: Don, any comments on that?

Member Yealy: Yes, this was one where the preliminary rating was high. Obviously, it's a composite of three different types of images and again, separate from all the other conversation we had, within the measure as it's created, there were really no voiced concerns about the composite construct. There are other concerns about the construct, but not this part.

Co-Chair Septimus: Okay, any other comments from Laura or Nancy, and then anyone else on the committee?

Do you see anything, Matt? I don't see anything.

Mr. Pickering: No, I don't see any chats or hands raised.

Co-Chair Septimus: Isaac, why don't you go to question number 7.

Mr. Sakyi: Voting is now open for measure 3621. The question is do you accept the Scientific Methods Panel's moderate rating for composite quality construct and rationale? The options are A for yes and B for no.

We're waiting for three more votes.

Voting is now closed for measure 3621. We have 100 percent vote to accept the Scientific Methods Panel's moderate rating for composite quality construct and rationale. The standing committee votes to accept the SMP's moderate rating.

Co-Chair Septimus: Very good. So I think the next

one is going to be feasibility, Don, right?

Member Yealy: That's correct. And again, the preliminary rating for this was moderate. That aligns with the type of measure that it is.

The data elements seem to be extractable in a fairly consistent and not terribly onerous fashion from electronic sources. It does require an installation though, this is not necessarily something that everyone can view from their base electronic platform. And right now there's participation fees regarding the American College of Radiology to submit the data, whether or not moving forward that remains a condition.

So overall, it does seem that the feasibility thresholds are met and it does not look to require either excessive investment or create either undue burden or irregularity in the actual collection of the information.

Co-Chair Septimus: Okay. Any comments then on feasibility?

Seeing none, I guess, Isaac, we can vote on that.

Member Yealy: This will be question 9 for everybody.

Co-Chair Septimus: Question 8.

Member Yealy: No, it's 9 since we skipped 8.

Co-Chair Septimus: It's 9, we missed it. You're correct. Feasibility, 9.

Mr. Sakyi: Voting is now open for measure 3621 on feasibility. The options are A for high, B for moderate, C for low, and D insufficient.

Voting is now closed for measure 3621 on

feasibility. There are 4 votes for high, 14 votes for moderate, zero for low, and zero insufficient. The measure passes on feasibility.

Co-Chair Septimus: Okay. Now we move into the usability and use. Obviously, this is not a maintenance measure.

Don?

Member Yealy: Yes, so you hit on the important point, Ed, in that it is not something that we're asking ourselves how is it gone, but oddly enough, there's information on this from the American College of Radiology. And this is actually in use under the MIPS performance criteria in the ACR registries.

So for measures that are brand new like this and not already part of either our previous recommendations or CMS, I'm actually impressed with the existing information we have about the potential use of this. And again, to me, it looks pretty straight forward.

Member Sood: So just to clarify, it's the potential for feedback and the fact that they are already giving feedback for a metric that's new, right?

Member Yealy: I don't know that it's for the potential for feedback, but is it possible for people to engage with it, is how I would interpret it.

Member Sood: Okay.

Member Yealy: And that looks like a resounding yes to me.

Member Sood: Thank you.

Member Yealy: And if it sounds very much like feasibility, they do kind of overlap. They're not the

same, but.

Co-Chair Septimus: But if this was a maintenance, it would be a must pass, we would have to see that it's being used. But this is obviously not a maintenance measure.

Member Yealy: Another shot across the bow for the stewards for four years from now.

Co-Chair Septimus: That's right. Comments.

Mr. Pickering: Hand raised. Hand raised by Yanling.

Co-Chair Septimus: Thank you. I can't see everybody. Yanling, I'm sorry.

Member Yu: No problem. I just have a question about use. It probably needs clarification. Is that only those who registered with American College of Radiology went into this type of measure or anyone can use this measure?

Member Yealy: So let me answer that. The data for usability came from the American College of Radiology, but moving forward and assuming that this -- I don't know that we'll eventually pass this, although we all know which way it's going. I don't think that's the predicate requirement for use of the measure moving forward from that. Just the data that are available to us were from the American College of Radiology.

Dr. Mahesh: Can I make a comment?

Co-Chair Septimus: Not necessarily, unless someone has a question directed to the developer.

In the chat, Sara, do you want to make your comment? Leapfrog hospital surveys use dose-related data, in case you haven't seen that.

Member Hawkins: Yes, there is information that's already used. For example, the Leapfrog surveys to rate facilities on their safety, patient safety, and are used in dose-related data.

Member Yealy: Right, so I just was uncertain if that was the measure as we're looking at it today. It's certainly a potential for overlap, but I wasn't certain. I don't know that part of the Leapfrog all that well.

And I'm not sure it's relevant to us. We know that the measure, as it's being presented to us, is not unfeasible, but actually used.

Member Hawkins: Absolutely. And that's in support of that.

Co-Chair Septimus: Okay, any other hands that I hope I haven't missed somebody again. Terry has to step away soon, but we'll probably take a break.

Okay, let's go, Isaac.

Mr. Sakyi: Voting is now open for measure 3621 on use. The options are A for pass, B, no pass.

Member Yealy: This is a question 10 if anybody's wondering.

Co-Chair Septimus: Right.

Mr. Sakyi: Waiting for one more vote. Is anyone having difficulty?

We have 18 votes. Voting is now closed for measure 3621 on use.

We have 18 votes, pass, and zero, no pass. The measure passes in use.

Co-Chair Septimus: That was a tough call, Isaac. Okay, usability.

Member Yealy: The, again, the preliminary rating for this was moderate. Not a surprise given the nature of the measure. And the data that we have existing comes from a voluntary, motivated group of reporters and showed a performance that's steady. For a measure like this, I'm not certain that a steady performance is exactly what we want because it might imply something else, but at the beginning of the journey where we're rolling this out more broadly, I think it actually suggests that there is opportunity still in usability and what that will be is I'm certain if we pass this it will be something we learn about four years from now.

Co-Chair Septimus: Comments from the committee?

Matt, any hands that I haven't seen?

Mr. Pickering: I do not see any.

Co-Chair Septimus: Okay, well, Isaac, let's go on usability.

Mr. Sakyi: Voting is now open for measure 3621 on usability. The options are A for high, B moderate, C low, and D, insufficient.

Member Yealy: It's question number 11, is that correct?

Co-Chair Septimus: That's correct. I must say this is an impressive committee. We've been at this for a while. We haven't had a break and people are still focused and asking great questions. So a real credit to all of you.

Mr. Sakyi: We're waiting for one more vote. Voting is now closed for measure 3621 on usability. We have 4 votes for high, 14 votes for moderate, zero for low, and zero insufficient.

The measure passes on usability.

Co-Chair Septimus: Okay, so the last is the final vote for suitability for endorsement. So Isaac will go to that next vote which should be number 12.

Mr. Sakyi: Voting is now open for measure 3621 on the overall suitability for endorsement. The options are A for yes and B for no. Waiting for one more.

Voting is now closed for measure 3621, the overall suitability for endorsement. We have 16 votes for yes, and 2 votes for no.

Co-Chair Septimus: Did you say it passes, Isaac?

Mr. Sakyi: With 16 votes for yes, the measure is therefore recommended for endorsement.

Co-Chair Septimus: I didn't want to steal your thunder.

Okay, we're about 30 minutes behind with the agenda stated. I really do think we need to take a break unless people feel otherwise.

Let's come back at 2:00 p.m. Eastern Time and resume with the next measure. Is that okay with everybody? Okay. See you back --

Mr. Pickering: So I just want to make sure -- sorry. That should be fine. Acumen is the next measure developer. Are you on the line?

Ms. Lin: This is Cheng from Acumen. It's okay that we can gather at two.

Mr. Pickering: Okay, great. All right, because your next two measures are up. so we'll reconvene at two. Thank you. And we'll see everyone back at two o'clock.

Member Falvey: Matt, I was actually going to be scheduled for another 2:00 to 2:30 meeting. If

Acumen is a developer for both the other measures. Is there any chance you can flip those two?

Mr. Pickering: You want to flip the two? I'm not sure what you mean.

Member Falvey: The falls measure, I have to go meet with and IH from 2:00 to 2:30. So if that one could be talked about second if it's not inconvenient.

Mr. Pickering: I think that's fine. We can start with 679, if that's okay with the developer. Maybe Cheng hopped off.

Ms. Lin: Yes, that would be fine.

Mr. Pickering: Thank you. Thanks, Jason.

Co-Chair Septimus: So 679 and then 674.

Mr. Pickering: Yes, that's it.

Co-Chair Septimus: Got it. Okay, see everyone. Bye-bye.

Mr. Pickering: Thanks, everyone.

(Whereupon, the above-entitled matter went off the record at 1:33 p.m. and resumed at 2:00 p.m.)

Mr. Pickering: Okay, so this is Matt, again, Matt Pickering, and just to see if everyone is coming back in. I have 2:00 p.m. on the eastern side, so I'm checking to see Ed.

Co-Chair Septimus: I'm here.

Mr. Pickering: All right.

Co-Chair Septimus: I'm back. Let's hope everybody else is coming back.

Mr. Pickering: I will say, Ed, that with Terry Fairbanks' message, he's going to be out until about 3 o'clock. And with Iona out and then we also heard from Jason, he'll be out for about 30 minutes. We should be down to 16 which is right at quorum, so we won't be able to really lose anybody else. If we do, we lose the quorum. We'll see how things go with coming back.

But I guess we'll get started. Before we do, I just want to check in real quick, too. It's now shortly after two. People are sort of coming back in. That's great. And is the developer, Acumen, are you back on line?

I see their name there, but I'm not sure if anyone from Acumen --

Ms. Lin: Hi, Matt. I'm on and I should have several other colleagues on the line with me.

Mr. Pickering: Okay. Great. Great, great, great. Excellent.

And just so Acumen is aware, we are proceeding with developer engagement through the meeting today. Unless the developer is asked for any clarification, we just kindly ask the developer to just to remain quiet as the standing committee deliberates. And then if there's questions where the standing committee would like the developer for clarification, we'll turn to you to do that, just to carry that forward.

Okay, Ed, I will go ahead and turn it back to you.

0679: Percent of High Risk Residents with Pressure Ulcers (Long Stay)

Co-Chair Septimus: I'm ready to go. As you named it, we're going to start with 0679 which is shown on the screen. This is a maintenance measure. I

believe it is also an outcome measure. It measures the percent of long-stay, high-risk residents in nursing homes who have Stage II to IV for unstageable pressure ulcers. And you can see the definition about defining long term as those who receive greater than 101 or more cumulative days of nursing home care by the end of the target assessment period. And you can see that the nursing home residents are defined as high risk if they're impaired, they're comatose, they're malnourished, or at risk of malnutrition. And this measure is based on the minimum data set that you can see there on the screen.

I also understand that some of these data sets are somewhat older, but we'll get into that discussion a little bit later.

So with that, I'll turn it over to the measure developer, who will give us a three to five minute summary. And then we'll turn it over to the primary discussant for our committee. And this is 0679, so it should be -- let me get this right.

Member Sood: Me.

Co-Chair Septimus: Geeta, right?

Member Sood: Right.

Co-Chair Septimus: I just want to make sure I get this right. Okay, measure developer, please, will you introduce yourself, please.

Ms. Lin: Thanks, Ed. This is Cheng from Acumen and thank you for the walk through of the design of this measure. We want to emphasize that pressure ulcer is important to address as they are one of the most unwanted and preventable adverse events.

Typically pressure ulcers occur in individuals with poor mobility, who experience sustained pressures

for long period of time. The occurrence of pressure ulcer can be effectively reduced with adequate staffing, staff composition, and proper staff education.

That said, this measure is based on MDS which is an assessment tool that is required to be built out on a regular basis. So there is no actual burden for nursing homes to collect measure information.

The measure has been endorsed several times. It was last endorsed in 2015 and this measure has a long history in the Nursing Home Quality Initiative which started in 2002. The measure is publicly reported through the Care Compare website and Provider Data Catalog. The Care Compare site, previously known as Nursing Home Compare, is a tool for patients and families to learn the quality of care at different facilities to select providers.

Provider Catalog, as the name suggests is a provider info database and is a site commonly used by providers and researchers.

In addition to use by the patient and research community, confidential feedback reports are available to providers through the CASPER reporting system. Providers can access patient level information for residents within the facility to identify areas for improvement and take action to work on pressure ulcer care and prevention.

The measure passed SMP for reliability and validity voting. Specifically, the data elements of the measure demonstrate pretty strong validity through the Red (phonetic) study. The result is still highly relevant because the relevant data elements in the MDS assessment forms hasn't changed since the study was conducted.

In terms of reliability, on top of the one-quarter

reliability that SMP rated as moderate, we have supplemented our final submission with our findings on the four-quarter reliability.

The four-quarter version is the main form of public reporting from the Compare site. CMS designed the rolling four quarter version of the measure to improve measure reliability, stability, and reportability. And indeed, we observed that the rolling four quarter version demonstrate much strong reliability compared to the one-quarter version.

When reviewing the committee worksheets, we noticed that there's one question that came up about the degree of variability between nurses staging pressure ulcers. We would like to clarify that CMS designed the measure to include stage II to IV and unstageable pressure ulcers. This measure does not differentiate the stages in measure calculation.

While it is possible that some nurses may stage pressure ulcer differently, the pressure ulcer will be captured by the measure regardless of staging as long as it's above one. This approach should mitigate concerns about the impact of inconsistency staging.

Okay, and this wraps up my intro. And like I said, I have several Acumen colleagues on the line with me and we thank the group for discussing and considering this measure.

Co-Chair Septimus: Thank you so much. So I will turn it over to Geeta who will take us through our elements starting with, on evidence.

Member Sood: Thank you. So thank you for that wonderful summary. And I will just say that I have large shoes to fill since Terry Fairbanks and Don Yealy did such a beautiful job with the first two

measures, so bear with me.

So as you nicely described the measure is a measure looking at the ulcers, Stage II through IV, and unstageable in long term care patients in this high-risk population, the impaired, bed mobility, comatose patients, and malnourished patients.

The data, as you also mentioned, is available through the minimum data sets. The other discussants that will be speaking about this measure with me are John, Joel, and Raquel. Hopefully, you are here and we'll start by talking a little bit about the evidence side which as I understand is evidence for the relationship between the outcome and the healthcare process. And that can be demonstrated by the prevalence, as well as the variability in performance.

So the data that was shared showed that there was a pretty significant prevalence of pressure ulcers in nursing home patients. And that the developer also shared a lot of data on guidelines and the evidence within those guidelines showing that measures are modifiable and can be -- those modifiable measures can, in fact, reduce the incidents of pressure ulcers for patients in nursing homes.

There wasn't a whole lot of committee comments outside of that for the evidence side.

Do my co-chairs or co-discussants have other things that they want to add for the evidence portion, evidence 1A portion?

Member Bundy: Geeta, I think you said it well. It just reaffirms what we already knew.

Member Sood: Thank you. Any other comments from the rest of the committee?

Member James: It's a very nice summary, Geeta.

Member Sood: Thank you.

Co-Chair Septimus: If John says that, that's really a high compliment, Geeta.

Member Sood: I appreciate that.

Member James: She's not done yet.

(Laughter.)

Member Sood: I was very impressed by the robustness of the evidence that the developer provided regarding not just the variability in performance, but actually tying the improvement, the measure -- the interventions that can be used to improve pressure ulcers to this measure as well. So I thought that was beautifully done.

So I'm a little bit newer to this. Ed, should we --

Co-Chair Septimus: So we're going to start with evidence. I think just we've sort of viewed that in your co-reviewers, I think of waiting on that. Now we just ask whether or not there are any questions from the rest of the committee.

Matt, can you monitor everybody? My screen doesn't show any hands.

Mr. Pickering: I don't see any hands up. Again, this is opportunity for questions or discussion related to evidence for this measure and I don't see anything in the chat.

Co-Chair Septimus: So let's go to Isaac, the master of the voting. So that should be obviously question number 1 of 11. Okay, Isaac?

Mr. Sakyi: Thanks, Ed. You should be able to see the measure we are voting right now, 0679, on your screen. If you're not able to see, you may need to

refresh your browser.

Co-Chair Septimus: I can see it in the voting application, but I can't see it on the current screen.

Mr. Sakyi: Yes, if you can see it in the voting application, that's okay.

Co-Chair Septimus: Do you want to read it, Isaac?

Mr. Sakyi: Yes. Voting is now open for measure 0679 on evidence. The options are A for pass, and B, do not pass.

Co-Chair Septimus: Now you've got it on. Good. So we are looking for 16 votes. Fifteen. So if there's anybody not able to vote -- okay, there we go. Fantastic.

Mr. Sakyi: We have 16 votes. Voting is now closed for measure 0679, evidence. We have 16 votes for pass and zero do not pass. The measure passes on evidence.

Co-Chair Septimus: Okay, Geeta, let's go to performance gap.

Member Sood: Thank you. The performance gap, the developer showed nice data showing that the median score was 8.6 and that there was a standard deviation of 5.1 percent and there were several facilities that had a score of zero. So there certainly does seem to be variability and, therefore, opportunity for improvement on the performance gap side.

I believe the performance gap side also includes disparities and the disparities were also measured for age, race, and socio-economic status. And interestingly, older patients above 85 had a lower rate of sacral decubitus ulcers. As expected, unfortunately, race, if you were non-White, you had

a higher rate of ulcers and if you had -- if you were of lower socio-economic status that also increased your rate.

So clearly there are differences by socio-economic status. As one of the committee comments mentioned, it's not really felt that those would be part of the pathogenic pathway and that we probably shouldn't adjust for those socio-economic factors because we wanted, in the interest of health equity.

One of the questions that I have for the developer related to this is I'm wondering if there is confounding with the age issue and the race issue as likely older patients are unfortunately more likely to be White than they are to be non-White. So I'm wondering if you have data on a multivariate analysis to control for that difference as well or what your thoughts are.

Mr. Pickering: Developer?

Ms. Lin: Geeta, yes. We can confirm there is indeed an interaction between race and age and indeed in the nursing home population that the White race population seems to be older and thus is indeed a contributor to the unexpected effects that we're seeing on age side.

And the other thing we can note is that because of the uniqueness of the nursing home population, the younger population are those not covered by Medicare and they are more likely to have some severe injuries like disabilities, maybe not qualify for the 30-month period yet. So that can contribute to the younger population staying in bed for a longer period of time.

Member Sood: Thank you for that answer.

Do my co-chairs, John, Joel, and Raquel, do you

have other thoughts?

Member Bundy: This is Joel, I had the same question about age and race. I was also thinking through about age and like diabetes, specific type 1 diabetes and maybe those patients were younger and they didn't survive to 85 and so you sort of self-selected or help your population that would be less likely to get pressure injuries. But Geeta, I think you asked the right question.

Co-Chair Septimus: Well, physiologically, Joel, they may have been a lot older than their actual chronological age.

Member Bundy: Yes.

Co-Chair Septimus: I'm beginning to ask the question, what is now considered older?

Member Bundy: Older than 80.

Member Sood: Definitely not anywhere below that.

Co-Chair Septimus: I'm going to be 75 next year, so I have to be very careful about how I define old.

Any other comments on performance gap from anyone?

Matt, do you see anyone?

Mr. Pickering: Yes, Yanling has her hand raised.

Co-Chair Septimus: Yanling, thank you.

Member Yu: Yes, thanks. I just am curious about the sponsor said there is non-Medicaid population shows higher risk for pressure ulcer than Medicaid population. I just want to know how they're related to socio-economic status when you talk about this type of a risk.

Any explanation of why they're part of a pattern?

Co-Chair Septimus: Are you asking the developer?

Member Yu: Yes.

Co-Chair Septimus: Okay.

Ms. Lin: So it seems that's there a question for us and we think it's really related to the uniqueness of different enrollment groups, and for example, when we look at the non-Medicaid population that includes one that Medicare only and this is a population that will be older. And so this is the part where the research shows older population may have lower function than others and also has the proportion of very long health care stays and the population may be sicker. And there's also a smaller population that's neither covered by Medicare nor by Medicaid. And this population, we just find in general to be more problematic and have some unique characteristics associated with them.

Member Sood: Thank you. For the developer, for the next round, would you be willing to consider preventing stratified rates as well? Certainly, I think our group doesn't think that risk adjustment would be appropriate, but potentially, it looks like you've done a fair amount of analytics on the different strata of socio-economic status, insurance coverage, and age. I think that would help us understand this issue a little bit better as well.

Ms. Lin: Yes, absolutely. This is on our to-do list for the next cycle and other measures.

Co-Chair Septimus: We'll so note that in our findings. Anything else on gap?

Matt, we okay?

Mr. Pickering: Yeah, I don't see any hands raised or

any questions in the chat box.

Co-Chair Septimus: Okay, Isaac, it's your turn.

Mr. Sakyi: Voting is now open for Measure 0679 on performance gap. The options are A for high, B for moderate, C for low, and D, insufficient.

We have 17 votes in. The voting is now closed for Measure 0679 on performance gap. We have ten votes for high, seven votes for moderate, zero for low, and zero insufficient. The measure passes on performance gap.

Co-Chair Septimus: Excellent. So, this is not a composite measure, so we don't have to go through that, and we can go right to scientific acceptability looking at reliability.

Member Sood: Thanks. So, for reliability, we want to be looking to make sure that the specifications of what the measure is measuring is in fact correlating to what we think is clinically important development, and that the testing is reliable in that one person will be measuring it the same way a second person will be.

So, I think the data that were presented were well presented. I do have a question for the developer related to a concern for surveillance bias, which was also brought up in one of the comments.

The numerator, from what I understand, is when there is documentation for a pressure ulcer. One could envision that a facility may not be so robust with doing surveillance or documenting surveillance of pressure ulcers and may conversely look like they're doing better than another facility that has very robust surveillance.

So, could you, developer, would you mind talking a little bit about your thoughts about this and how

well, what you think in terms of capturing that?

Ms. Lin: I will try to answer, and if the other team from Acumen has more clinical knowledge, please feel to jump in.

I think one part that can relieve some of the concern there is that this measure is looking at stage two to four and unstageable pressure ulcers, and from stage two, the wound is really open, so I think that will make the wound easier to detect compared to a closed injury.

So, that's just a part that I can think of, that by focusing on more severe pressure ulcers, that will mitigate some concerns of surveillance issues.

Member Sood: Okay, thank you. I will talk briefly about some of the other reliability factors and then ask my co-discussants and the rest of the committee to comment on this portion if that's okay.

So, some of the analytics that you did with RAND were quite nice, looking at a gold standard nurse, their evaluation of pressure ulcers with a facility nurse, and that was found to be very highly correlated between the two, which I think speaks both to reliability and potentially validity since the gold standard nurse could be considered a gold standard.

The only issue with those data are that those data were 13 years old as they were presented.

There were other data presented regarding changes over time, and that seemed to have a slightly bit of a concern in that the number of pressure ulcers, thankfully, is a low volume outcome, and there was some variability in terms of whether a facility performed or changed from one decile, two deciles, or three deciles over time.

There was about 33 percent of facilities that changed more than two deciles over time.

So, the other reliability test that was done was the split half test, and that showed good reliability, about 0.33 and 0.50 for the signal to noise in terms of the reliability of this metric.

Those were my summary of the reliability portion. Co-chairs, John, Joel, and Raquel, comments, thoughts, or other questions to add?

Member Bundy: Geeta, this is Joel again. I'll make one comment on the first questions about surveillance bias.

Having previously worked in a long-term care facility and currently working in acute care, I do recognize that stage two and above, if you look, you can see it, but that means you have to look, and poor performing facilities may not look as often as others, so I still think that's an issue.

I don't think that's something we can solve today, but that was certainly something that struck home with me.

Member Sood: Thank you. That's very helpful to have your insight on that. Other thoughts or comments from either co-discussants or the rest of the group?

Co-Chair Septimus: Anything, Matt?

Mr. Pickering: I don't see any hands raised or questions in the chat box.

Co-Chair Septimus: All right, I think you're on.

Member Sood: So, before we vote, if I could just ask the developer also, the issue of the old data that was presented for reliability. For the next

round, if you could please have more current data, I think that would be helpful as well.

Ms. Lin: Yes, something I wanted to add for the RAND study is that they did do a follow-up study four years after they did the original one. Basically, things are consistent and they find the results should be very similar to what they had done before.

In looking at the actual MDS form, we see that the item construction, like the design of the item is the same on the MDS form compared to when it was first designed. So, I think that based on this finding, the RAND study results should be still very relevant at this point.

Member Sood: Thank you.

Co-Chair Septimus: Yeah, that was my comment when I introduced it for both of these measures, that the data is somewhat old, but you're saying, if I understand it, that you've got some subsequent stuff that says it's still valid?

Ms. Lin: Yes.

Co-Chair Septimus: Okay, all right, okay, Geeta, is it okay if we go ahead and go forward? Okay, okay, you're on again, Isaac.

Mr. Sakyi: Voting is now open for Measure 0679 and the question is do you accept the scientific methods' panel's moderate rating for reliability? The options are A for yes and B for no.

Mr. Pickering: And just a reminder, SMP looked at both of these measures, and in this case for 0679, passed the measure on reliability.

Member James: That really helps. Is this five?

Co-Chair Septimus: You mean what number is this? This is number three.

Member Roney: Three.

Co-Chair Septimus: This is number three, John.

Member James: Okay, all right, okay.

Mr. Sakyi: We have exactly 17 votes. Voting is now closed for Measure 0679. We have 17 votes for yes, zero for now. The standing committee votes to uphold the SMP's moderate rating for reliability.

Co-Chair Septimus: So, now we go to validity, Geeta?

Member Sood: Thank you. So, for validity, we're looking to make sure that the data elements correctly identify the actual data that we want to use, and that the measure actually correctly identifies differences of quality of care for what we're looking for.

The way that validity was measured in this metric was by comparing it to other quality metrics, including the star rating for long-term care facilities. That's not my favorite way of measuring and evaluating validity, but that is one that I understand is often most convenient and sometimes the best that we can do, but nevertheless, that's the way that the validity was measured, and Matt, I think, will add that the scientific method panel considered that to be a moderate level rating for validity.

The correlation coefficients were good when you compared this metric with other quality metrics related to long-term care. So, co-chairs, anything else to add, John, Joel?

Member Bundy: Nothing here.

Co-Chair Septimus: John?

Member James: Yes, so I'm not quite sure where to insert this, but I think maybe this is the place. I don't want to take on the scientific panel, but I want to think about the future, and maybe it's how I would use this information.

So, let's say that seven and a half percent of the people have a two to four-stage ulcer in a given facility, so how many of those are four, which to me shows a lot more patient neglect than some twos? And a stage four pressure ulcer is much more. It has risk of infection and a number of other really adverse outcomes.

So, in the future, I would like the sponsor of this to look at the proportion of twos, threes, and fours, and unstageable pressure ulcers that they're finding in these facilities.

I think that may provide some interesting insight, and it certainly gives me as a patient looking at this information a better feel for how good the nursing home is doing with this stuff.

Co-Chair Septimus: So, yeah, thank you, John. I think the developers said in the beginning that they're all lumped together. They're not broken out.

Member James: Yeah, and if they were broken out, I think that would help a lot. It might be very interesting actually, I think.

Co-Chair Septimus: Right.

Member James: And I don't know how practical that is. My sense is that it probably wouldn't be hard to do.

Member Sood: My understanding from what the developer said was that there was a reduction in

reliability when you do that because there is variability between what one nurse will stage in one place versus another nurse in that same --

Co-Chair Septimus: Yeah, I think your kappa scores are going to go down for that.

Member James: Right.

Co-Chair Septimus: But that is so interoperator variable, especially if you're talking about twos and threes. I mean, fours, I think most people, I hope, can spot a four.

Member Bundy: And John, this is Joel. We can, you know, like in a hospital, that's easy to do. You can, you know, do a two, three, or four because you have the same wound care ostomy nurse sort of roaming through the entire building, so then you have great reliability and validity, but I can see here where that would be more difficult across the country.

Co-Chair Septimus: Well, I think we'll take that as a recommendation, John, and thank you for that. Any other comments about validity?

Member Sood: So, I just wanted --

Co-Chair Septimus: Elissa has one.

Member Sood: Oh.

Member Charbonneau: Yeah, I was just going to comment that in the inpatient rehab sector, we have to report on new or worsening pressure ulcers, which is really, I think, a better reflection of how we are preventing, because pressure ulcers, you know, once you have one, you're seeing the tip of the iceberg as we all learned, and, you know, that chances are that it's worse than it looks.

So, I'm just wondering if there was any thought to looking at new or worsening pressure ulcers instead of just the static measure?

Co-Chair Septimus: Developer?

Ms. Lin: Yeah, I'm happy to jump in here. First, thank you very much, John, for the recommendation of breaking it down by the different stages. That's something we're definitely interested to do and supporting our next round of maintenance or even recertification.

And for the point about the new or worsening pressure ulcers, this is something that there is an existing NSQIP quality measure like specifically focused on new or worsening since admission.

And the reason that we're not doing this for the nursing home population is that when you look at the population base, these are people who have stayed in a nursing home for at least 100 days.

So, as we compare it to status change since admission, it's really like looking for a very persistent pressure ulcer, and I think it should be within the facility's responsibility to care for that case.

Co-Chair Septimus: Yeah, I think that's an important change. Obviously in the acute care hospital, there are hacks where if it's not present on admission and they get a stage three or four, that constitutes a hack.

Here you have people that are in nursing homes for three or more months, so I think it does become a little bit more difficult. Does that help, Elissa? I can tell. If you're not happy, I'm not happy.

Member Charbonneau: You know, I think another thing that we haven't really discussed is that in

patients who are in long-term care for over 100 days, those are patients that, first of all, Medicare doesn't pay for that, so that's a very large Medicaid population.

So, usually patients have to spend down so they qualify for Medicaid to get coverage to be in those facilities, and they're obviously very impaired from a functional standpoint and require a high level of nursing care.

So, I think, you know, having worked in those kinds of facilities with so-called low level patients, that it really is very important to make sure that their pressure ulcers are not worsening, because that's when they develop sepsis, as we talked about earlier today, or other significant complications, so that's just perhaps my bias.

Member Sood: That's -- I did not know that and thank you for sharing that, but actually that is also a perfect segue to the threats to validity and risk adjustment. I forgot is that a separate question or is that part of the -- is there just one question for validity and we talk about the risk adjustment now?

Mr. Pickering: Great question, Geeta. So, within the validity assessment, risk adjustment is part of that, as well as exclusions to the measure, et cetera. These are all considered threats, and risk adjustment is trying to control for any potential bias.

With that actually, Geeta, going into that, I did want to recognize Theresa Edelstein as she had made a point around looking at pressure ulcers in terms of their setting or origin.

Many patients are admitted from hospitals to nursing facilities with pressure injuries and vice versa at different levels of severity.

So, when thinking about the risk adjustment piece, Geeta, I'm unsure if the developer has looked at place of origin or setting of origin for pressure ulcers, and maybe was thinking is that something in the risk adjustment model, but I will turn it over to you, Geeta, to see if you want to answer that question.

Member Sood: Thank you.

Co-Chair Septimus: I was just about to call on Theresa. Okay, keep going.

Mr. Pickering: I'm sorry, Ed.

Co-Chair Septimus: That's okay. No, it's great. Go ahead, Geeta.

Member Sood: Theresa, did you want to comment verbally or to add to what Matt said?

Member Edelstein: Thank you, Geeta. My only additional comment is I've worked for many years in a long-term care as an administrator and I've run pressure ulcer collaboratives as part of my work at the New Jersey Hospital Association, and I would really like to see a measure that looks at the patient and is patient centric, not setting centric.

Pressure ulcers, pressure injuries are something we can all have an impact upon if we work together across care settings, and I think the sooner we get to that, the better for the patient. So, that's just my additional comment and perhaps my bias.

Member Sood: Thank you. It's great to have expertise on the panel so that we can learn from the things that you know, and hopefully that will also help the developer refine their metrics as well, so thank you both for sharing your thoughts about this.

So, if it's okay, I'll move briefly to the risk adjustment or the other aspects of validity before we vote. Does that sound okay?

Co-Chair Septimus: Yeah, please.

Member Sood: Okay, excellent. So, there was some -- the threats to validity really are to make sure that there are -- there's no bias, that there is risk adjustment as needed, and whether, like Matt said, whether inclusions and exclusions are appropriate.

So, as we talked about before, this is considered a high-risk patient population, which would be patients that have impaired mobility, are comatose, or are malnourished.

To my understanding, it doesn't look like the developer is risk adjusting for pressure ulcers, and there was some thought about, in the comments, about whether that's a good idea or a bad idea.

I personally have somewhat mixed feelings about this because I think you could make the argument that this is, quote, unquote, a never event, and that as such, all of these pressure ulcers should be preventable and therefore risk adjustment may not be quite as important to determine an expected amount of infection.

On the other hand, we also know that patients are very varied, as many of you have taught me, in terms of their risk for pressure ulcer development.

So, co-discussants, do you have other thoughts about validity and risk adjustment, and then obviously the rest of the committee?

Member Bundy: I do not.

Co-Chair Septimus: Okay, rest of the committee?

Member Charbonneau: You know, I'll just add that, you know, if you can stratify looking at pressure ulcers by diagnosis or patient population, so there's a lot of differences.

So, someone who is paraplegic or quadriplegic in a wheelchair, you know, of course has a much higher risk of developing a pressure injury than someone who is more mobile, and then of course the elderly frail with their skin conditions are also at very high risk.

And it's just such a complicated topic just in terms of how patients are moved, and what they're sitting in, and their nutrition, and, you know, it's just a very complex issue, and I almost feel like this measure is oversimplifying.

And also by lumping everyone together, unstageable and -- you know, I'm not a wound care nurse, but I'm a physician, but, you know, lumping together stage two, three, four, and unstageable all together, I'm not really sure that that is the greatest measure, and then excluding stage one where that's where you have the opportunity to prevent somebody really from developing a worse pressure injury.

So, I don't know if anyone has any other thoughts on that, but I just almost feel like this is too simplistic of a measure.

Member Sood: I think that's such a great idea, collecting data on stratified or otherwise on different groups to help better understand what they should risk adjust for if we decide to do that.

And I see in the comments that Don doesn't think that all of them are preventable and neither does Theresa or David.

So, and the way that you're describing the fact that

it's a simple measure, I'm interpreting to mean that you would be in favor of risk adjustment or more refined patient information for the metric.

Co-Chair Septimus: I think, you know, having worked somewhat in this space, I don't know if we're letting perfect be the enemy of good here, but if it becomes too complex and there's too much variability in how people capture this and they make it too burdensome, is that going to impair progress?

I don't know the answer to that, but the measure developer may want to respond to that and see, mentioning some of the suggestions that committee members very, very well stated. What's the feasibility? We're not on feasibility yet.

What is the practicality of some of these suggestions in your world?

Ms. Lin: Yeah, thank you so much for giving me the opportunity to respond. And we do see that there is like a tradeoff between simplicity or robustness versus a patient-centered approach through risk adjustment, and we're open to that.

Something I wanted to bring up is that the underlying MDS instrument is going through some changes in the next couple of years, and currently some of the states, they're already collecting like new information that will impact the specification of this measure.

So this gives us the perfect opportunity to revisit this measure holistically, so this is on our plan and all of these great commendations of looking at stratifications, risk adjustments is something CMS and us are working towards to better prepare for the respecification of the measure in the near future.

And the reason that for those runs we are not

proposing any specification changes is that this measure, this robust approach has been proven to be working, and we see that there is differentiation in provider performance, and it's capturing the good performers and relatively worse performers.

But this redesign is definitely, like we're on the same page with you. We're definitely interested to look into the different stages, risk adjustment, and potentially including more information from the patient history before or outside of nursing home facilities.

Member Sood: Thank you.

Mr. Nagavarapu: And this is Sri from Acumen. I was just going to add real quickly that, you know, some of the types of information that you all are bringing up, I think the Section GG transition that Cheng is talking about is going to give us an opportunity to maybe get more consistent data on these sorts of things and make it more possible to look at them in a consistent way, and so it's something we're excited about.

Member Sood: Thank you.

Co-Chair Septimus: So, certainly, you know, those suggestions will be captured. Any other comments to finish your section on validity, Geeta?

Member Sood: No, I would just say thank you for those comments. It's hard to prove causality in terms of actually improving outcomes, but I think, you know, as we learn more, we can refine and make things better, and it sounds like you as the developer are interested in at least looking at big buckets for risk stratification that would, I think, help us in the next round as well, so please keep us updated and let us know how we can help with that.

Co-Chair Septimus: Any other comments?

Member Yu: Yeah, could I make a comment? This is Yanling Yu.

Co-Chair Septimus: Of course, Yanling, please.

Member Yu: Yeah, I understand this measure also is really a complicated issue based on the age, mobility, and other status, but this measure is for patients who are long-term care residents that have been in the facility for over 100 days.

So, if you have one stage one pressure ulcer, and when you're coming in, you evaluate where the patient is in nutritional status and mobility status, and then you can take, you know, a proactive measure to try to prevent deterioration of the pressure ulcer involved into stage two and stage four.

So, I think there's a role where this measure would help the facility to really, at the beginning, to watch all the status and, you know, then try to prevent it during this 100 days, you know, deterioration of the pressure ulcer.

So, I think no risk adjustment, I think, is reasonable. When I make a comment, I think no patients or residents would be, you know, excited to have deterioration of a pressure ulcer in those settings.

Co-Chair Septimus: Any other comments? I think we've had a great discussion on this, so I think it's, Isaac, I think it's time to vote.

Mr. Sakyi: Voting is now open for Measure 0679. The question is do you accept the scientific methods' panel's moderate rating for validity? The options are A for yes and B for no.

Member Sood: And this is question five, right?

Co-Chair Septimus: It's question five.

Member Sood: Okay.

Mr. Pickering: Jason, I see that you're back. Is that right?

Co-Chair Septimus: Yes, Jason's back.

Mr. Pickering: Okay.

Member Falvey: Yeah, I'm here.

Mr. Pickering: Thanks, Jason. Are you adding your vote in? There you go, okay.

Member Falvey: Yeah, I voted.

Mr. Sakyi: Voting is now closed for Measure 0679. We have 89 percent voting for yes and 11 percent voting for no.

Co-Chair Septimus: Okay, so we are now going to go on to feasibility.

Mr. Sakyi: The standing committee votes to uphold the SMP's --

Co-Chair Septimus: I'm sorry.

Mr. Sakyi: -- rating for validity.

Co-Chair Septimus: Sorry, Isaac.

Mr. Pickering: Yeah, we just have to capture that on record.

Co-Chair Septimus: No, no, no, I'm sorry. I meant I should know this by now. Okay, Geeta, feasibility?

Member Sood: Ye, so the next few should be relatively easy since this is a method already in use. Feasibility is really to make sure that this is not burdensome to collect the data to be able to collect

this measure.

So, I am not an expert in this area, but my understanding is that from what I have learned is that the data sources used, the MDS, is very readily available and not burdensome to obtain these data to be able to assess for decubitus ulcers.

My co-chairs who would know more about this area than me, any comments or thoughts?

Member Bundy: No.

Member Sood: Anybody from the rest of the committee?

Co-Chair Septimus: Well, Isaac, we just may have to go to vote on it. The preliminary rating on this was high.

Mr. Pickering: Elissa, you still have your hand raised. I just wanted to make sure you didn't want to have -- okay, you put it down. All right, sorry, go ahead, Ed.

Co-Chair Septimus: Now we're going to vote on feasibility.

Mr. Sakyi: Voting is now open for Measure 0679.

Co-Chair Septimus: That's question number seven, I believe.

Mr. Sakyi: The options are A for high, B for moderate, C for low, and D, insufficient.

We're waiting for two more votes.

Mr. Pickering: Has anyone not voted yet for feasibility? I see three A. We're still missing two votes. Anybody having issues? There is one. We're missing one vote now.

Member Yu: I voted wrong question, so I just corrected myself.

Mr. Pickering: We got it now. Thanks, Yanling.

Mr. Sakyi: Voting is now closed for Measure 0679, feasibility. We have 13 votes for high, five votes for moderate, zero for low, and zero insufficient. The measure passes on feasibility.

Co-Chair Septimus: Okay, and the next one is use. This is a must-pass criteria. This is a maintenance measure. Geeta?

Member Sood: So, this also I think is relatively easy. Use should be accountability and transparency. In other words, the data are available to the facilities and publicly, and that the facilities and the public have the opportunity to offer feedback, and that that feedback is then used to improve the metric.

So, these data are publicly reported, so I think that pretty much satisfies the accountability piece, and it sounds like there are opportunities through CASPER reports, et cetera, for the facilities to be able to offer feedback back to the developer.

Any other comments from John, Joel, or Raquel?

Member Bundy: None.

Member James: No.

Co-Chair Septimus: Anybody from the committee? This is on use. Again, it's a must pass. It's a maintenance measure. Anything, Matt?

Mr. Pickering: I don't see any in chat or the hand raised, so we're good.

Co-Chair Septimus: All right, Isaac, you're on.

Mr. Sakyi: Voting is now open for Measure 0679 on use. The options are A, pass, and B, no pass.

Co-Chair Septimus: And this is question number --

Mr. Pickering: Eight.

Co-Chair Septimus: -- eight. I just wanted to make sure, okay, just to capture transparency, okay.

Mr. Pickering: It should say 4a1. It looks like we have 18.

Co-Chair Septimus: Right, I just wanted to make sure.

Mr. Sakyi: Voting is now closed for Measure 0679 on use. We have 18 votes for pass and zero no pass. The measure passes on use.

Co-Chair Septimus: All right, so the next to last one is usability, and Geeta?

Member Sood: So, usability is to make sure that the audiences can use these data for accountability and performance improvement activities, both for improvement and to benefit the community and the population at risk.

It does appear that the percentage of pressure ulcers have decreased slowly over time, so, and these data are publicly reported, so both of those show that it is part of an accountability plan for usability.

Co-chairs, anything that you wanted to add in addition?

Member Bundy: Geeta, this is Joel. I would have liked to have seen more improvement than what was already seen. I'll take what I can get, but it would be nice to see better.

Co-Chair Septimus: Well, in fact, Joel, that you mentioned this, in my reading of this, that the scores may have slightly decreased over the last two or three years. Am I reading this correctly, developer?

Ms. Lin: But that's good, right?

Co-Chair Septimus: So, I think that addresses your comment, Joel.

Member Bundy: Yeah.

Co-Chair Septimus: Any explanation for that, developer?

Ms. Lin: Sorry, Ed, I missed the question.

Co-Chair Septimus: Well, the scores, as I read it, the scores slightly decreased, and the question is, is there any explanation why they may have decreased? Did something change or -- this is --

(Simultaneous speaking.)

Co-Chair Septimus: I call this BC.

Ms. Lin: -- decreased very slightly.

Co-Chair Septimus: Before COVID.

Ms. Lin: Yeah, it is. It has decreased very slightly and I think part of the decreases you may show in a seasonality trend as something related to seasonality, but when you look at kind of moving average, it's really a pretty stable measure with minimal decrease over time, and that can be a sign of potential improvement among the provider community.

Co-Chair Septimus: Okay, any comments from -- John?

Member James: Yeah, just a quick reminder to the presenter or the sponsor, it would be much more usable for me as a patient if I knew which level or stage of pressure ulcer was being looked at over the spectrum, so something for the future.

Co-Chair Septimus: Thanks, John.

Member Sood: The other -- thank you. The other comment that I wanted to make, which is probably not for the developer, but how this measure is used is that, of course, post-COVID with all of the changes that we've had, whatever baseline we had for all of these kinds of quality metrics is likely going to change.

So, I don't think that it's really the developer who is setting the value-based purchasing part or the specific numbers around that, but I wanted to just bring that up as well.

Co-Chair Septimus: Thank you, Geeta. Terry is back, right, Terry?

Member Fairbanks: I am, yeah.

Co-Chair Septimus: Good, did you have a good stiff drink when you were gone there?

Obviously, you haven't heard all of this discussion, so I'm assuming, Matt, that Terry would not vote on this last element anyway?

Mr. Pickering: It depends on where he's come into the conversation, but if he's just back in without hearing any of the usability conversation --

Co-Chair Septimus: I think he just -- you just came back in, right, Terry?

Member Fairbanks: Yeah, I'll hold. I'll omit any vote now.

Mr. Pickering: Okay.

Co-Chair Septimus: So, just to let you know. Okay, any other comments before we vote? Okay, Isaac, you're on.

Mr. Sakyi: Voting is now open for Measure 0679 on usability. The options are A for high, B, moderate, C, low, and D, insufficient.

We're expecting 18 votes.

Co-Chair Septimus: All right, who is the holdout?

Mr. Pickering: Anybody have any challenges? It should be, 4b1 is what it should say for usability. Last call, anybody having any issues? This is question nine. There we go.

Co-Chair Septimus: There's the 18.

Mr. Sakyi: Voting is now closed for Measure 0679 on usability. We have four votes for high, 12 votes for moderate, two votes for low, and zero insufficient. The measure passes on usability.

Co-Chair Septimus: Excellent, see, I waited that time, Isaac, I waited.

Mr. Sakyi: I appreciate that.

Co-Chair Septimus: Okay, so I think the last one is going to be overall suitability for endorsement?

Mr. Sakyi: That is correct.

Co-Chair Septimus: Okay, so why don't we let everyone go ahead and vote?

Mr. Sakyi: Voting is now open for Measure 0679 on the overall suitability for endorsement. The options are A for yes and B for no, and now we're expecting 19 votes.

We're waiting for one vote.

Mr. Pickering: I think Terry was going to hold off on voting.

Co-Chair Septimus: Yeah, Terry wasn't going to vote, so 18 is the right number.

Mr. Sakyi: Voting is now closed for Measure 0679 on the overall suitability for endorsement. We have 17 votes for yes and one vote for no. With 17 votes for yes, the measure is therefore recommended for endorsement.

Co-Chair Septimus: Okay, Isaac, thanks. And by the way, Geeta, you did terrific. Boy, the discussants, the primary discussants today have been outstanding, so you're really setting the bar high, and Theresa, I don't want you to get nervous.

So, we're going back, if I understand, to Measure 0674, correct?

Mr. Pickering: That's correct.

Co-Chair Septimus: Okay, do you want to put on the next slide?

Mr. Pickering: And I'll just mention I know that on our agenda, we have a break at 3:00 p.m., but since we are a little behind, if it's okay with the standing committee, if we just continue with 0674, then we'll sort of reassess where we are when we get through with that measure depending on the time we have for the remaining portion of the day.

0674: Percent of Residents Experiencing One or More Falls with Major Injury (Long Stay)

Co-Chair Septimus: Matt beat me to it. Well, if you look at the agenda, even though it said 2:00, we're 30 minutes behind, and we were going to think

about taking a break after these two measures, so, but thank you.

Okay, you can see it's 0674, percent of residents experiencing one or more falls with major injury in long stay. This is another CMS Acumen measure.

This measure reports the percentage of long-stay residents in nursing homes who have experienced one or more falls resulting in a major injury.

And this is important, how it's defined is fractures, joint dislocations, closed head injury, altered mental status, and subdural hematoma reported in the look-back period no more than 275 days prior to the target assessment, and then long term is defined the same way receiving more than 101 cumulative days, and again, this is also reported through the minimal data set.

So, with that, Theresa is the lead discussant with Jason as the secondary, and Arpana, I hope I said that correctly, as the secondary, so I'll turn it over to Theresa to talk about the evidence.

Member Edelstein: Okay, thank you, Ed. Does the developer need to make a statement this morning or --

Co-Chair Septimus: Oh, I'm sorry. I'm sorry. Thank you. Thank you. The developer has three to five minutes, forgive me.

Ms. Lin: No worries. I thought we were saving time, but we'll keep it short because there's a lot of similarities between the two measures.

So, first we want to emphasize that falls with major injury has high impact on health outcomes. We find it to be an important measure because research has demonstrated that injurious falls are the leading cause of disability and death for nursing home

residents.

Research has also shown that the proper living accommodations, amenities, and sufficient support by appropriate qualified staff can mitigate fall risk and prevent or reduce falls.

The measure uses MDS, and so there is no extra burden collecting information. Same as the pressure ulcer measure, the falls measure is part of the nursing home quality initiative. There are a lot of similarities between these two measures in terms of long history and wide use among the patient and provider community.

The measure has been endorsed several times and was last endorsed in 2015. The two measures share the same public reporting platform, namely the Care Compare website and the provider data catalog.

Confidential feedback of the cross measure is also available to the provider. Providers can access patient level information to identify areas of improvement and take actions for fall prevention.

Again, the RAND study on MDS shows that the measure has strong data element validity. Our testing results demonstrate that the one-quarter version of the measure has moderate reliability and the four-quarter version, which I should emphasize is the main form of public reporting, has high reliability.

Specifically as shown in the appendix we submitted, the four-quarter split half ITC is 0.63, signal to noise median score is as high as 0.8. Since the four-quarter version is the main form of public reporting, when we ran it through the SMP, SMP advised us to use the four-quarter version results for future maintenance, so that's just one part I wanted to clarify regarding the reliability testing and this

wraps up my intro.

Co-Chair Septimus: Thank you very much. Sorry that I didn't let you speak. Thank you for the people reminding me. Anyway, now, Theresa?

Member Edelstein: Okay, thanks, Ed, and having Geeta go first with 0679 really kind of laid the groundwork for this measure.

So, just a quick couple of comments on the updated evidence that was provided, there are some new sources of literature in the developer's submission that point to structural and process interventions that facilities can make to have a significantly positive impact on residents and the prevention of falls with major injury.

In the interest of time, I won't go through in any detail what they presented, but I found it to be certainly compelling and updated, and supportive of the continuation of measuring falls with major injury in long-term care facilities for long-term residents.

So Jason or Arpana, do you have anything you want to add?

Member Falvey: I mean, nothing too striking. I think the evidence that they submitted was also very strong, you know, not just for patient-level factors that we can address with these, you know, facility-wide interventions, but there was a lot of like facility-specific characteristics in terms of staffing and in terms of, you know, reducing uses of restraints and things like that that were really organizational and market level, you know, that I think the developer did a nice job bringing back and kind of showing that this is very modifiable at a facility level and even if your patients are relatively complex.

Co-Chair Septimus: Okay, anybody else on the

committee have a comment? Okay, well then I guess we will then vote on evidence. As you know by now, it's a must-pass criteria, and Isaac, you will tell us what to do. So, again, go forward. It's a new measure. Go for it, Isaac.

Mr. Sakyi: Okay, you should be seeing 0674, the title of the measure on the screen on the voting platform.

Co-Chair Septimus: The voting platform, yes, but not -- okay.

Mr. Sakyi: On the first question, voting is now open for Measure 0674 on evidence. The options are A for pass and B, do not pass. We have 16 votes in.

Mr. Pickering: Does anyone have any issues voting? So, after you see that title screen, you'll have to remember you have to hit the next to go to the importance measure report 1a. So, we have 16. There's 17. We should have two more votes. We're confirming, Terry, that you're voting this time?

Member Fairbanks: I am. I voted on question one. Are we just doing both in a row?

Co-Chair Septimus: No, we're just doing the first one.

Member Fairbanks: Yeah, I did that.

Mr. Pickering: Okay, is anyone else having issues voting? There's 18. Last call? Okay, so with 18 votes, that's over quorum.

Mr. Sakyi: Voting is closed for Measure 0674 --

Co-Chair Septimus: Because theoretically, we expected 19, right?

Mr. Sakyi: That's correct.

Mr. Pickering: That's correct.

Co-Chair Septimus: Okay.

Mr. Sakyi: So, with 18, we're past quorum. Voting is now closed for Measure 0674 on evidence. We have 18 votes for pass and zero do not pass. The measure passes on evidence.

Co-Chair Septimus: Thank you, Isaac. Okay, Theresa, gap?

Member Edelstein: Okay, so over time, there has been some slight improvement in this measure. On page 35 of the developer's submission, they noted that slight improvement.

There is still room for improvement with some facilities scoring a zero, which is a good score to have in this instance. It makes for -- 19 percent of facilities with a rate of zero makes for some issues with reliability, which we will get to, but overall, there is still room for improvement in this measure.

There were some differences discussed related to age, race, and socioeconomic status, but there was no -- well, there was some discussion about whether risk adjustment is warranted on this measure and perhaps we could discuss that.

Co-Chair Septimus: So that seems the theme across both of these measures. Okay, any other comments from the committee on gap?

Member Falvey: I mean, this is Jason. I will just point out that there was some, you know, disparities across socioeconomic and race. The race disparities were somewhat counterintuitive.

Generally, patients that are minoritized typically in nursing homes have higher rates of complications and there's a lot of evidence to suggest that in other

outcome measures.

But I was wondering if the developer wanted to comment on, you know, if they disentangled that a little bit or what they thought that gap was, because it was pretty substantial and I didn't really see much reason that they would see that.

Co-Chair Septimus: Developer?

Ms. Lin: Yeah, for this, we find that race could be something related to staffing levels and that's why like one side there's maybe before these sides, that where there's a higher rate of minority races, there is lower staffing level, and we know that staffing level is highly relevant in this measure and can cause a higher rate there.

Member Falvey: I'm sorry, maybe I'm reading this incorrectly, but your measure states that your white only population has almost double or more than double the risk for falls than non-white populations, so your evidence just suggested the opposite, so that's why I'm maybe a little bit confused.

Co-Chair Septimus: Developer? Can you repeat that, Jason? Just repeat it again.

Member Falvey: Absolutely, I'm sorry. Maybe I'll just, I'll put a page number, too, so you can see where I'm reading from. So, I'm reading from the measure worksheet and it's on page three where it's the summary.

And race, it says the developer reports that the white only population, 3.73 percent, is at higher risk for experiencing falls with major injury than the non-white population, 1.8 percent.

And my general understanding is facilities that serve minoritized residents are exactly where -- you know, what you said was that the staffing levels are

often worse. Minorities disproportionately have to use lower-quality nursing homes with less star ratings.

So, you know, I would have expected higher fall rates, and so I wonder if it's something with the reporting. Are they less likely to have falls reported, which I'll bring up when it's appropriate, but that's kind of my thought.

Mr. Nagavarapu: Thanks for raising that point. This is Sri from Acumen. This is certainly something that we could look into more.

I suspect, as you do, that there is an interaction with other effects going on, including like location of nursing homes, as well as enrollment composition, so whether that's through, let's say, Medicare or other payers, and so this is definitely --

I agree with you that result is somewhat counterintuitive given other patterns that we typically see in nursing homes, and it's something where I think what we could do is look at interactive effects with other information.

I do think that the upcoming information that will be available in the new Section GG will allow us to do a more detailed job of that type of analysis, just in terms of the type of information it will provide on functional status and so on, if part of what we're seeing here is differences across race and ethnicity in either true functional status or like reporting levels across nursing homes, but, yeah, this is definitely an area that we'd be interested in looking more into down the road.

Member Falvey: Okay, well, that make sense. Your thought is it's an interaction or confounding at some level of, you know, other patient characteristics versus facilities that have a higher percentage of

non-white residents or minoritized residents are less likely to be reporting falls, you know, masking or being poor reporters of that information, which would, you know, one would be more concerning than the other clearly.

Member Charbonneau: I'm just wondering, you know, GG is already being collected, so do you have any thoughts as to when you would start to use the GG data? It's not really that new anymore.

Mr. Nagavarapu: Yep, no, that's true. It looks like the states have started collecting GG data as of last October, and so for a measure like this that has a relatively long look-back period across assessments for long-stay residents, and so we'll need the data to build us up a little bit from that period. But we're hoping that as soon as we start to get a few quarters of data that can build up with GG, that we'll be able to do a more nuanced look at some of these questions.

Co-Chair Septimus: So, is there any other concerns about -- again, I think we're on gap, right?

Mr. Pickering: Right.

Co-Chair Septimus: Okay, I just want to keep people focused, so is there any other questions just around the gap? I think we all admit that gap still exists. There is some nuances about how that might be looked at and addressed, but that there is a gap. It looks like there is a gap, but does the committee need any further discussion? Okay, let's vote. Go, Isaac.

Mr. Sakyi: Voting is now open for Measure 0674 on performance gap. The options are A for high, B, moderate, C, low, and D, insufficient.

Mr. Pickering: We're missing a few votes.

Mr. Sakyi: Voting is now closed for Measure 0674 on performance gap. We have one vote for high, 17 votes for moderate, zero for low, and zero insufficient. The measure passes on performance gap.

Co-Chair Septimus: Thank you, Isaac. The next one is reliability, and just to remind everyone, this was reviewed by the SMP. Theresa?

Member Edelstein: Okay, so a lot of the issues that were discussed under 0679 pertain here as well, in terms of the age of the data. I think the developer addressed those concerns in our prior discussion.

There was a kappa for gold standard to gold standard nurse and the kappa was 0.967. The gold standard to facility nurse agreement was 0.945.

The average signal to noise reliability score was 0.45, but as I said before, with 19 percent of facilities achieving a perfect score of zero, this results in lower variation of scores between facilities.

So, there were some concerns by some of the SMP members that this was low, but appears to have been explained by the developer.

Co-Chair Septimus: Jason?

Member Falvey: No, I think that was a good description, I think, in terms of, you know, comparison to nurse, you know, to nurse reviewers. I think this is, you know, pretty highly reliable.

Co-Chair Septimus: Okay, committee response or questions?

Mr. Pickering: Just to confirm, Elissa, do you have a -- your hand is raised. I just wanted to make sure --

Member Charbonneau: Yeah, I mean, my only comment is that it's very hard to believe that any facility would have a zero reporting of a fall with injury over this period of time.

Co-Chair Septimus: Elissa, you mean it's hard to believe they have a zero event, but I could believe almost any report.

Member Yealy: Actually, I looked at that and I said zero, really? That's concerning.

Co-Chair Septimus: I've also never been wrong, Elissa, not once, so.

Member Roney: Elissa, you read my mind exactly. I'm sitting here going why am I such a pessimist about this. Maybe they could have been that good. It's just hard to believe when we see as many falls as we do in acute care.

Co-Chair Septimus: Well, remember that's falls with injuries, so it's probably -- I don't want to say the tip of iceberg, but there are probably many, many falls that don't result in injuries.

Member Charbonneau: It's still, you know, knowing this population, it's very, very hard to believe. And I don't know if there's any way to audit that, but it's just really inconceivable to me.

Member Falvey: Well, yeah, I have a comment. But I mean, it fits better in validity, so we might hold that and come back to it, because that's going to come up.

Member Edelstein: I would also say that, you know, there's been a lot of work done in nursing facilities based on best practices and evidence-based literature over the decades that I've been involved in this field. So I mean, 19% of 16,000 nursing homes is not that many.

And it is possible that there are that many facilities that have incorporated those best practices and actually have avoided major injury, at least the ones described in this measure. That's not to say there aren't other major injuries that occur as a result of falls. So I understand the skepticism, I'm just offering a different perspective.

Co-Chair Septimus: Thanks, Theresa. Any -- does anyone else have their hands raised, Matt?

Mr. Pickering: I don't see any, and I don't see any in the chat box.

Co-Chair Septimus: And I don't see anything new in the chat box either. Okay, so let's vote on reliability, Isaac.

Mr. Sakyi: Voting is now open for Measure 0674. The question is can you accept the SMP's moderate rating for reliability. The options are A for yes and B for no.

Co-Chair Septimus: Question Number 3.

Mr. Sakyi: Waiting for everyone to vote.

Co-Chair Septimus: Here you go.

Mr. Sakyi: Voting is now closed for Measure 0674 on accepting the SMP's moderate rating for reliability. We have 94% saying yes, with 6% saying no. The standing committee votes to uphold the SMP's moderating rating for reliability.

Co-Chair Septimus: Okay, thank you, Isaac. And then Theresa, we will now go to validity, which is another must-pass criteria.

Member Edelstein: Okay. So, similar to the measure we just talked about, the performance score validity was established by correlating with other measures

of nursing home quality and the five-star ratings.

These were low -- there were low but positive correlations between facility performance and other quality measures. Almost all of them fell below 0.1.

Co-Chair Septimus: Jason, you said you had some comments on this?

Member Falvey: Yeah, so I don't want to, you know, start a lengthy conversation about this. I think, you know, I think my major concern is going to be we're capturing falls based on self-report, which the facilities control. So MDS is filled out by the nurses that are employed by the facility, so there is some concern about bias there.

And a recent article that came out in Health Services Research, which is a very good, you know, methods-type journal, and it came out, I'm looking at the year right now. Recently, 2022, so very recently, it compared MDS with Medicare claims verification of fall-related trauma. So for all of these, there's a pretty specific ICD-10 codes associated with fracture, closed head injury with altered consciousness.

And MDS reporting only captured 57% of all falls in the facilities, with a pretty wide gap in terms of minorities. So minoritized patients are much more likely to have falls not reported on MDS than White patients. And that gap, let me get it so I have it accurate.

So they identified 150,000 falls and claims, 57.5% were reported. Long stay was better. But White patients had 60% of falls reported versus non-White at 46%. And that was pretty consistent even after adjusting for some facility-level characteristics.

So I'm happy to send that out, but I think it raises the broader question where pressure ulcers are not

coded well in hospital data. And I don't think, you know, we should be thinking about claims for that.

But on a measure like fall-related traumas, reasonably easy to capture in claims, and Medicare does claims verify, you know, other hospitalization and ED-type of events for home care and other -- and other post-acute care settings.

So maybe the developer can comment on why they haven't kind of shifted this towards a claims verification, or at least done some sort of double check to see how significant the under-reporting is.

Mr. Nagavarapu: This is Sri from Acumen. Thanks a lot of the question. Yes, this is paper that we're aware of also. And actually it's part of the motivation for some monitoring that we're planning to do.

So essentially for all of the post-acute case measures, including the long-stay measures, we conduct quarterly monitoring and annual monitoring for CMS as part of our work with CMS supporting them on this.

And what this paper, as well as other sort of ongoing questions about assessment coding in the MDS has brought up, is this exact question that you're talking about, is the extent to which there's under-coding and how it varies across different types of facilities.

And so our plan, and you know, this hasn't been incorporated into the testing materials that you have today, but our plan is with the quarterly monitoring to essentially link information to Medicare claims data, as you said.

And here, with the specific measure we're talking about today, we can focus on the specific sets of major injuries that are included in the numerator

and map those to the correct diagnosis codes in Medicare claims. And look at the rates, just as you're saying, right, and compare the rates for different types of facilities, different types of subgroups and so on.

I think that's something we're planning to do. There is an inherent limitation in that in the sense that we'll be able to do that for the Medicare population in a very effective way.

For the Medicaid population, that's Medicaid only, that becomes more dicey to the extent that there is a information on falls in T-MSIS, which is the Medicaid federal information system. We can pick it up, but reporting quality in T-MSIS can be dicey.

And then the other limitation is for Medicare Advantage patients. Their inpatient claims and potentially outpatient claims related to falls with major injury are not consistently reported encounter data that CMS has.

But having said that, I think the exercise that you're suggesting is still valuable. You would just be constrained to looking at those involved in Medicare fee-for-service to really get a good sense. And it's something that we're planning to do.

Member Falvey: I think that's helpful. I mean, I, you know, I get those limitations. Working with that type of claims data pretty frequently for research, I do understand that there's some quality issue there.

You know, there is all, you know, state all-payer claims databases and things that these things would be adjudicated against in terms of, you know, state by state. But it's going to be highly variable.

But I think I would like to see that number in the future to, you know, to know if there is some serious biases that we have to consider in reporting.

Mr. Nagavarapu: That definitely makes sense. An all-payers claim database suggestion is a really good. Like, I guess in principle we could -- you could try to get access to and use OSHPD, for instance, in California.

And so we could look into all-payer databases where you might not have something across the whole country, but could have something for a specific state that could give us a window to this.

Co-Chair Septimus: Jason, thanks for the comment. Just to let you know, look at the chat box, there's a couple of links there that have been placed that some people may want to read. Are there any other comments on validity? Great discussions.

Okay, well, I guess any hands up, Matt?

Mr. Pickering: Yeah, I don't see any hands raised and no additional questions in the chat box, Ed.

Co-Chair Septimus: Okay, well then we will have Isaac do his magic.

Mr. Sakyi: Voting is now open for Measure 0674. And the question is do you accept the SMP's moderate rating for validity. The options are A for yes and B for no.

Co-Chair Septimus: So that's question number 5, correct?

Mr. Pickering: Yes, that should be 5.

Co-Chair Septimus: You skipped, I just want to make sure people skipped.

Mr. Pickering: Right, right.

Mr. Sakyi: We have 17 votes in so far.

Mr. Pickering: Anyone having issues, challenges?

Okay.

Co-Chair Septimus: Got the 18 -- 19, hey.

Mr. Pickering: All right, there's the straggler. Okay.

Mr. Sakyi: Voting is now closed for Measure 0674 on accepting the SMP's moderate rating for validity. We have 19 votes for yes and zero for no. The standing committee votes to uphold the SMP's moderate rating for validity.

Co-Chair Septimus: Okay. The next one, Theresa, is going to be feasibility. Theresa? Feasibility.

Member Edelstein: Okay, can you all hear me okay? Because I was having some bandwidth issues. Okay.

Co-Chair Septimus: Now you're back.

Member Edelstein: Okay. So feasibility very similar to the last measure. All the data elements are defined fields in the nursing home MDS, collected electronically and accessible electronically.

Mr. Pickering: So Elissa, do we have a disclosure of interest for our new committee member on the screen?

Member Charbonneau: No, he is very quiet.

Co-Chair Septimus: Jason, you have a follow-up?

Member Falvey: No, no, I think, yeah, I agree with what Theresa said.

Co-Chair Septimus: Okay, anybody else, Matt, have their hand up? You can see the whole thing, I only see half of it.

Member Falvey: No, yeah, there's no one has their hand up and I don't see any questions in the chat

box.

Co-Chair Septimus: Okay, well, if none then I think we will vote then on feasibility. Isaac.

Mr. Sakyi: Voting is now open for Measure 0674, on feasibility. The options are A for high, B moderate, C low, and D insufficient.

Co-Chair Septimus: Well, we got 14.

Mr. Pickering: So again, this is Question 7. So say -- it should say Feasibility 3A.

Co-Chair Septimus: You guys want a break? Fifteen. Here we go.

Mr. Pickering: Still missing a few. Anybody having any challenges? Again, this is a --

(Simultaneous speaking.)

Co-Chair Septimus: You guys are in Washington, we've heard a lot about voter suppression. Is that what we're seeing here, Matt?

Member Roney: I was sure worried about them in Washington at one point. My goodness, I had to check on them.

Co-Chair Septimus: Still?

Mr. Pickering: So let's see --

Co-Chair Septimus: I think it's ongoing.

Mr. Pickering: Okay, so it's --

Co-Chair Septimus: We're still missing two or three.

Mr. Pickering: So we're at 16 votes. Anybody having any issues? Again, this is feasibility 3A. This would be slide 7. Any issues with votes? There's another

one. So we had 19 last question, now we're down to 17. And I know that Joel, you're --

Member Roney: I think I might have voted on the wrong one, Matt. Hold on.

Mr. Pickering: Okay, sure, thanks, Jamie. And I know -- Joel, did you vote?

Member Bundy: I did.

Mr. Pickering: Okay, because I know you're stepping away at four.

Member Bundy: Yes.

Mr. Pickering: Okay. So there's 18. Jamie, was that you, did you just vote?

Member Roney: Yeah, it was me. I voted on 6 instead of 7. I'm sorry I was off again, oops.

Co-Chair Septimus: No, it's easy to do that. No, no, no, I did the same thing a couple times ago. It's easy. So we're up to 18.

Mr. Pickering: Okay, last call for Feasibility 3A before we close it. Okay. Isaac, let's close it.

Mr. Sakyi: Voting is now closed for Measure 0674 on feasibility.

Mr. Pickering: There's -- now we're at 19.

Mr. Sakyi: We have 19. We have seven votes for high, 12 votes for moderate, zero for low, zero insufficient. The measure passes on feasibility.

Co-Chair Septimus: Thank you, Isaac. And then we go to use, which is a must-pass criteria since this is a maintenance measure. So Theresa, anything to?

Member Edelstein: This measure is publically

reported in Care Compare, it's used in accountability programs. Facilities have access to their data through CASPER, same as -- same as the previous measure.

Member Falvey: Nothing to add. Nothing to add, us researchers are very thankful as publically reported and in databases so we can study it too, so.

Co-Chair Septimus: Okay, Matt, anybody else have their hand up?

Mr. Pickering: I do not see anyone having their hand raised.

Co-Chair Septimus: Okay, well, let's vote on use.

Mr. Sakyi: Voting is now open for Measure 0674 on use. The options are A for pass, B no pass.

Co-Chair Septimus: And is that going to be 4A1?

Mr. Pickering: Yup, that's correct.

Co-Chair Septimus: Okay, wanted to make sure that everyone's on the same page.

Mr. Pickering: So after the feasibility question you should just hit next to go this question, 4A1 on use.

Mr. Sakyi: We have 17 votes so far.

Mr. Pickering: There's 18. And anyone still not voted for that 19th vote? And this is 4A1, use. Okay, last call. Okay, Isaac, you can just close it.

Mr. Sakyi: Voting is now closed for Measure 0674 on use. We have 18 votes for pass and zero for no pass. The measure passes on use.

Co-Chair Septimus: Okay, then usability.

Member Edelstein: Okay, so as I mentioned earlier,

there has been a slight improvement over time in this measure. So it has usability, as evidenced by the continued incremental change that we see in a positive direction.

Co-Chair Septimus: So unlike the other measures, this one's improved, slightly. Jason.

Member Falvey: Nothing to add there. Just, you know, very low percentage, so even incremental improvements are good.

Co-Chair Septimus: Other comments from the committee? Matt, anything?

Mr. Pickering: I don't see any hands raised or questions in the chat box.

Co-Chair Septimus: Let's go to usability.

Mr. Sakyi: Voting is now open for Measure 0674 on usability. The options are A for high, B moderate, C low, and D insufficient.

Co-Chair Septimus: This is Question No. 9.

Mr. Sakyi: We have 17 votes, 18. Voting is now closed for Measure 0674 on usability. We have five votes for high, 13 votes for moderate, zero for low, and zero insufficient. The measure passes on usability.

Co-Chair Septimus: Okay, that last one of course is suitability for endorsement. Isaac.

Mr. Sakyi: Voting is now open for Measure 0674 on the overall suitability for endorsement. The options are A for yes and B for no.

We have 18 votes. Voting is now closed for Measure 0674 on the overall suitability for endorsement. We had 19 votes, we have 19 votes for yes and zero for

no. The measures is therefore recommended for endorsement.

Co-Chair Septimus: Excellent. So let me ask the committee, we're scheduled to take a break. We can shorten that break, if -- gets a quick cup of coffee or take a bio break, instead of 15 minutes come back in ten. It's up to the committee what it wants to do.

Mr. Pickering: Bio break. Okay, head nods, head nods.

Co-Chair Septimus: All right, two -- excuse me, 3:50 back start. So take a quick break, a little bit shorter than was scheduled. Thank you.

(Whereupon, the above-entitled matter went off the record at 3:41 p.m. and resumed at 3:51 p.m.)

3501e: Hospital Harm - Opioid-Related Adverse Events

Co-Chair Septimus: Okay, why don't we go ahead and get started. We're certainly going to get through one measure, we may not get through both of them. We definitely want to make sure we have time for public comment because we probably will have public comment today. So we want to be respectful to people who are waiting to publicly comment at the end of a long day.

Okay, so the next measure is an e-measure. It's a new measure, 3501e, a hospital harm opioid-related adverse events. The measure steward is IMPAQ International.

It's the measure that assesses the proportion of inpatient hospital encounters where patients ages 18 and older have been administered an opioid medication and subsequently suffered the harm of an opioid-related adverse events, as administered

by an opioid antagonist with 12 hours. This measure excludes opioid antagonist administration occurring in the operating room setting.

So with that, the lead discussant is going to be Yanling Yu, but first we'll turn it over to the measure developer, if they can give us a three to five minute summary of this measure.

Dr. Hall: I sure will. This is -- my name is Kendall Hall from IMPAQ, an affiliate of the American Institutes of Research, and I'm joined on the call today by Chana West and Bo Feng from our ECQM testing team. And we're pleased to present Measure No. 20 -- 3501e, hospital harm opioid-related adverse events.

Opioids are important part of in-hospital pain management, but with opioid prescribing comes the potential for adverse outcomes, including altered levels of consciousness, respiratory depression, and the potential for hypoxic or anoxic brain injury or even death.

This facility-level electronic clinical quality measure, or ECQM, assesses the proportion of inpatient hospital encounters during which patients age 18 years or older have been administered an opioid medication and are administered naloxone, an opioid antagonist, within 12 hours, as an indicator of an opioid-related adverse event.

This measure excludes naloxone administration occurring in the operating room setting.

As many of you may recall, we brought this measure to the committee in the spring of 2019 and have since revised the specifications using the excellent feedback we received at that time. The denominator population has been now limited to encounters where patients received at least one

opioid during the hospitalization.

We've also addressed the need to better establish the connection between the naloxone in administration and the opioid-related event by including a requirement for any naloxone administration to be preceded by an opioid administration by hospital staff during the encounter to count as a harm.

We have also included a single time parameter for that administration to count as a harm.

The measure value sets have been updated to include all opioids and to ensure that the most current codes are used and that the codes harmonize across other ECQMs in current CMS quality reporting programs.

Lastly, we remain steadfast in our assertion that there is enough variability in the measure score across sites to demonstrate a gap in care. Testing data showed that across our six hospital test sites, harm rate in some sites can be as large as four times the harm rate in other sites.

The absolute number of harm events may be small compared to other harms, but when they do occur, opioid-related adverse events have been found to increase hospital length of stays, healthcare costs, 30 days -- 30-day re-admissions, and as I mentioned, the inpatient mortality.

So the relativity matters and the variability indicates ample room for improvement. With the revised specifications and supported by our testing of feasibility in 23 hospitals and implementation testing in six hospitals, we believe that we have addressed the concerns raised by our original submission.

So thank you for allowing our team to present this

measure to the committee. We're grateful for your consideration for endorsement, and we look forward to addressing your comments and questions.

Co-Chair Septimus: Excellent, succinct review. Yanling is the lead discussant, secondary will be Ed and Sara. And just to remind the committee, David Stockwell is a recusal.

Go for it, Yanling, lead us through the first with --

Member Yu: With evidence?

Co-Chair Septimus: Yeah, evidence, why not.

Member Yu: Thank you, Ed. I mean, it's going to be hard to follow those previous performance before us, but we will do the best we can.

The sponsor, and Ed I'll point out, this is a -- remember, this is a new measure, and it's an outcome measure. And the -- to support the evidence of this measure, the sponsor cited a number of references published from 2009, 2020, including a report by Joint Commission, IHI, CMS.

And this report shows how naloxone administration has been used by the adverse drug event surveillance as a indicator of a severe opioid-related adverse event.

So the sponsor believes that by encouraging hospital to implement evidence based on the practice, such as routine patient monitoring for patients at adverse effect of opioid, the measure can lead to better quality care by reducing excessive opioid administration and the use of naloxone in hospital setting.

So the preliminary reading for evidence is past. And the overall, the committee's evaluation on evidence is supportive of -- and ranging from appropriate to

strong.

So Ed and Sara, do you have anything to add on the evidence? I know, Ed, you shared an article about it. I don't know where you want to bring this up.

Member Pollak: Well, I think maybe somebody can help me with where I should put this, because I'm happy to -- so I had questions about the connection between harm from respiratory depression and outcome and opioid dependence and inappropriate opioid prescribing in the inpatient setting, which we all agree exists.

And whether the outcome in this measure, which as I understand it is naloxone administration, whether there -- after opioid administration, is tightly linked in terms of the evidence.

I saw some citations from the Joint Commission database, which I actually oversaw the last three years when I was patient safety officer. And I know that that would be based, for example, on their self-reported sentinel events, which is a tiny subset of all events.

So it's not that I doubt that there's strong evidence that respiratory depression is a significant harm. It's the connection of naloxone use as a surrogate for that, for those harm events. And I don't know where to fit that it.

Co-Chair Septimus: I think -- I think, Ed, that that obviously this is an outcome measure, you want the evidence to link to an outcome. So I think -- I think it's reasonable to discuss this under evidence, unless Matt thinks otherwise.

Mr. Pickering: No, this is an appropriate place to --

Co-Chair Septimus: I think this is an appropriate time.

Mr. Pickering: Especially if you're talking about that evidence linkage, Ed, and that's where you're concerned about as well. This would be a good place to discuss that. And if the committee doesn't have anything to add to that, maybe the developed can provide any clarification to your questions, Ed. But I'll let the committee respond first.

Co-Chair Septimus: Yeah, that's Ed P, not Ed S. Any comments from the committee?

Member Yealy: So I have the same basic concern, and I think I shared it the last time that you reviewed this. And particularly so when the antagonist administration within 12 hours is linked as a harm. As I wrote in the chat box, I think we're going to catch a lot of dolphins in this tuna net.

Co-Chair Septimus: Okay, got it. Yeah, I think I remember you -- that was when you first got on the committee, Don.

Member Pollak: So the article I shared was just an example, one of many, and maybe just for the measure developer to understand. Outside of emergent naloxone administration, there's also evidence that it's not only acceptable practice but good practice to administer low-dose naloxone infusions perioperatively, which goes beyond the operating room use of it.

And so if you're talking about, you know, big bolus doses of naloxone, then I would agree. But if you're talking about small doses, those are often used to prevent pruritus and bowel dysfunction, as well as other things, and a little different. So how will we -- I guess I'm wondering what the evidence is that this measure will capture that distinction as well.

Co-Chair Septimus: You're asking this to the measure developer, Ed?

Member Pollak: Well, I'm asking it to the group as well as to the measure developer. I want to put that idea out there that I'm not sure that there's a clear way to distinguish that.

Member Yealy: Ed, it's Don, I'd like to link onto that. Naloxone is often given for both diagnostic and therapeutic reasons. So it's often given with a modest to no effect, yet I'm not sure the measure can distinguish from said -- people who get funny, either in the emergency department inpatient setting, it's often throw a few things at them, particularly -- naloxone's fairly harmless if you're wrong.

Yet we'd be linking that to a very specific assessment. And I see this every day. I would actually tell you three-quarters of the naloxone use I see ends up not being indicated.

Member Yu: I think under the reliability discussion by the Scientific Method Panel, there's one member raised a question about how the measure can differentiate between the use of naloxone as a indicator or -- of opioid-related adverse events, versus other use follow, or in combination with other opioid use. So I think that that's a valid question.

Co-Chair Septimus: So Geeta has written something in the chat box. Geeta, you want to ask the developer your question?

Member Sood: Sure, it was actually more of a suggestion, because one of the great things about being on this committee is that I learn so many things, and I did not know there were so many different indications for naloxone.

I wonder if it would be valuable to do a chart review of different -- in different hospitals of patients who

were administered naloxone to better understand if that evidence is fair or not, that it is an indication of harm versus all these other indications. So just a suggestion.

Ms. West: Hi, can you hear me?

Co-Chair Septimus: It's used as an antibiotic all the time, right? I think it's a pretty good question, so can the developer sort of respond to some of the comments?

Dr. Hall: Yes, sure, so -- I'm sorry.

Ms. West: Oh, I'm sorry, this is Chana. Kendall, we were both chiming in to say the same thing.

Dr. Hall: Sure.

Ms. West: I just wanted to speak to the testing that we actually did. When we did our testing, we actually did ask that nurse reviewers take a look at the indications that were provided for the patients that did receive naloxone. So they looked at POS scores, so, you know, standardized evaluation tools. Looked at the response, not only the indication but also the response that was provided.

And a great number of those cases did provide -- basically the patients were unarousable and had some respiratory challenges, and that's why they administered it. So we did look at that.

And then in terms of the response for that, most of them said that after receiving it that the patients were more awake. There were some instances where they didn't actually specify it, but the great number of them actually indicated that it was given for opioid reversal and for issues with arousability.

Member Pollak: So I don't -- thank you very much. I don't think Don or I was suggesting that it's not the

most common reason it's given. So in a large sort of sample, to Don's point, we would expect that most commonly, it would be given for opioid overdose. But it's not the only reason, and I'm wondering how we can exclude.

You know, if you're a, you know, a place that has no cases of opioid-related adverse events and you're a relatively small facility, how will this measure enable you to not -- and you do decide that the best thing is to adopt one of these protocols for your spine patients undergoing scoliosis surgery or hysterectomy patients to avoid giving too many opioids. How will you distinguish that?

And in a large, national survey I get that it'll be hard to not see -- all that will wash out. But for the individuals, I'm struggling to see how appropriate care will not be -- individual facilities, how appropriate care will not be misclassified as inappropriate care.

Dr. Romano: Could I add, this is Patrick Romano, also from the measure development team at IMPAQ International. Could I just add a comment?

Co-Chair Septimus: Patrick, you want to follow up in response as the measure developer, then?

Dr. Romano: Yes, if I could --

Co-Chair Septimus: Okay, then please do, yeah, please.

Dr. Romano: Supporting my colleague, Dr. Hall. So I just wanted to call your attention, as Chana pointed out, this hypothesis was very specifically tested through our field testing, as described in the measure score validity section of our testing submission. So very specifically, the nurse abstractors went through to see how many of these events were opioid-related adverse events.

And the accuracy was basically 98%, the positive value was 98% at one site, 100% at the other sites that were included in the testing. So this I think provides a pretty solid reassurance that at least at the hospitals that were involved in our testing here, we weren't picking up anything that wasn't related to opioid administration.

At -- now, of course we understand the point that in some cases these may not be very serious or a full-blown respiratory failure where somebody ends up back on a ventilator. But still, we also looked, as you can also see in the testing report, we looked at how often the nurses documented that there was an improvement in the patient's level of alertness.

And so we found that the majority of the time, the medication was being given not because the patient was in imminent respiratory failure but because they were cognitively depressed. And the nurses in fact documented 76% of the time that the patients became more awake after the naloxone administration.

So this is why we think that naloxone administration in this particular setting within 12 hours of opioid administration, intravenous systemic opioid administration, seems to be a good marker of an opioid-related adverse event.

Mr. Pickering: Can I make a quick comment?

Co-Chair Septimus: Of course, I was going to ask you if you wanted to comment.

Mr. Pickering: Sorry, if I can -- Ed, if I can just interject, I apologize. And I'll give you a comment. It sounds like we're starting to get a little bit into the validity portion, and I think your comment, Ed, initially was talking about maybe evidence to support this naloxone use within this, and kind of.

So I think we started getting into the evidence discussion, now we're getting into the testing and testing information and validity as Dr. Romano has also shared some of the results. So we just want to keep that in mind, that we're really talking about evidence.

And with any outcome measure, right, part of that evidence is best -- is to see if there's associations that the actual accountable entity in this case can do like a structure or process that the accountable entity in this case can implement to improve the outcome that's within the measure. And so does the evidence support that and does it align with the logic model that the developer has mentioned.

I appreciate the questions here because, you know, it is a good question around is there evidence to support this naloxone use the measure within the measure.

But if we're getting into some of the testing and the actual validity of that, keep in mind that that's starting to get into the validity assessment. And maybe some of those comments and questions could be -- hold off on those until we get to validity discussion.

Co-Chair Septimus: Matt, I was thinking, you know, this is a -- I agree with your comment. I thought this was a gray area, because I think that -- I think we as a committee have to sort of figure out the link that naloxone -- this is an outcome measure.

And so I think we sort of have to make sure that this marker and measuring it is going to lead to a better outcome. And if there wasn't a correlation, then that -- then that, then the evidence becomes pretty thin.

But Ed, why don't you finish up your comments. But

I agree, I don't want to get too far ahead, but I thought there was too -- so much of an overlap here with the evidence, I thought we had to discuss it. But thank you, though.

Member Pollak: I think this comments that I'll make next is connected to the evidence. So I actually after seeing that they had rates of 100% and so on of appropriate, essentially appropriate naloxone administration, the literature to see what the connection between naloxone administration and indicated administration of naloxone was.

Because as Don suggested, I've clinically seen it always used, as he mentioned, as throwing a very wide net and giving it to a lot of people during, for example, respiratory RRT and code situations as protocols, so many of those people would have had a stroke and it wouldn't actually be addressing -- or it wouldn't actually be addressing inappropriate opioid use.

So when I saw rates of 100%, I was -- and it doesn't mean those people didn't get opioids at some point within the last 12 hours, it just wasn't the root cause of what was going on. So I struggle with that disconnect of the published literature on the evidence, because I didn't see rates anywhere near 100% of connection of harm and respiratory arrest, respiratory -- opioid-related respiratory arrest with naloxone administration.

Co-Chair Septimus: Okay. Great discussion. Anything else that we need to discuss on, again, we're talking about the evidence, okay. The extent to which the measure focuses is evidence-based and is important to making significant gains in quality and outcomes. Just to keep focused on that definition for evidence. Any other comment?

Mr. Pickering: I don't see any hands raised at this

point, Ed.

Co-Chair Septimus: Okay, well, let's go vote then, Isaac.

Mr. Sakyi: You should be able to see on your screen the measure number and the title. If not, please refresh your browser. Voting is now open for Measure 3501e on evidence. The actions are A for pass and B not pass. We are expecting 16 votes.

Co-Chair Septimus: Did you say 1-5?

Mr. Sakyi: One-six.

Co-Chair Septimus: Oh, 1-6, excuse me. Thank you. There you go, there it is.

Mr. Pickering: Sixteen.

Mr. Sakyi: Voting is now closed for Measure 3501e on evidence.

(Simultaneous speaking.)

Mr. Sakyi: -- pass and six do not pass. With ten votes for pass, the measure passes on evidence.

Co-Chair Septimus: What's the percent?

Mr. Pickering: It has to be more than 60, so --

Co-Chair Septimus: So what's the percent? So we don't make a mistake. Sixty percent?

Mr. Pickering: 62.5%.

Co-Chair Septimus: Okay, good, so I just wanted to make sure that that does meet the 60-plus percent mark.

Okay, all right, so the next one then is going to be gap. So I'll turn it back over to Yanling.

Member Yu: Yeah, as far as the performance gap, it was tested in six hospitals in five different states with very bad site. And the range in performance across six hospital was from .11% to .45%. The overall performance rate about six hospital was .34%, with a standard deviation of .12%.

The preliminary rating for opportunities for improvement is moderate. And the performance also stratified for disparity by age, sex, race, ethnicity, that no risk adjustment -- no risk stratification was applied to this measure.

And the committee members' comments pretty much diverged. The two members rate the perform gap as low citing very small performance range from between 11 and 40 -- .11 and .45%. I'm wondering if the measure developer had any thoughts on whether this measure is robust enough to capture some very small number of events. Other noted a moderate performance gap.

Co-Chair Septimus: Sara, are you -- Sara, you had a comment in the chat box you might want to.

Member Hawkins: Well, as we move into now this gap here, this is a concern I have, is that again, the only exemption that's been made so far is an OR setting. But as has been discussed, there are so many variabilities into why someone may be receiving opioids and where their risk is.

And as we consider risk stress assessment, you know, considering even the opioid-naive patients versus not, and again, how that's going to affect, you know, the specificity of what we're able to measure is just a concern that I've had as I've read through these initial -- the evidence piece and to this point.

Co-Chair Septimus: Well, this is the gap. Is there a

gap in care? Ed, did you have a comment?

Member Pollak: What I saw was, and I couldn't really -- the numbers were fairly low, and I'm doing it from memory, which is always a bad thing. But it was like .12 to .46. Somebody who -- I should defer to somebody who's actually looking at it.

Co-Chair Septimus: That's pretty good. It's .11 to .45. That's pretty good.

Member Pollak: Oh my gosh, my failed memory. It worked for once. So anyway, I guess I don't know and I would just ask the committee. Is that good or bad? That's like I, you know, tell my team, anytime you give a number, you need to say what the -- and you know, I would maybe defer to Sara and Yanling.

I don't -- I don't know what you all think. Is that -- and other members of the committee, is that an acceptable gap or not acceptable gap? I don't know if that's?

Member Yu: The standard deviation is .12%. So I just wondering what is the noise level for that one.

Dr. Hall: Is that a question for us?

Member Yu: Yeah, yeah, what is the noise level?

Dr. Hall: Yeah, I'll send it to Bo to talk about the signal to noise.

Dr. Feng: Hello, everyone, can you all hear me?

Co-Chair Septimus: Yeah, Bo, it's fine.

Dr. Feng: Thanks very much. First of all, thank you very much, Yanling and Ed and everybody else for the great comments.

So, I have heard a couple questions. Starting from the first one, based on memory, question number

one is about the performance gap. So that is essentially, when I interpreted it, the committee's concern about whether or not the measure is topped out.

So keep in mind that, so first of all, I recognize that the measure rates at their base value or in the absolute value that appear low. But we shouldn't be really focused on their absolute values. Instead, we should pay more attention to their relative values.

Look at across all the six test sites that participated in our measure testing, the harm rate in some sites can be as large as four times of the harm rate in other sites. That's point number one in terms of the variability and the performance gap. To me, that indicates ample room for improvement.

Number two is about the robustness. So I am an applied health economist. When I'm thinking about the question's robustness, I'm always thinking about if I tweak the little things on the margin, to what extent my results hold. To what extent they are stable. You know, we can definitely, certainly test the robustness, particularly among the small-sized community hospital.

As you all know that this called empirical question, the first step, and it's usually a really key step, is to do data collection. And again, at the -- in my view, measure implementation can definitely help.

So just to reiterate, we have six test sites participated in our measure testing. And all these six test sites are not homogenous. Some are really large hospitals located in urban areas, while others are small-sized community hospitals located in rural area. And again, the harm rate ranged from .11% to .45%. Again, relatively, in my view, that matters.

Co-Chair Septimus: Okay, so again, we're -- again,

we want to focus on gap. Does anybody else have any comments about gap? Is there a gap in care?

Member Pollak: I mean, I'll defer to others on this group, but I think it depends what you mean. From, I think that Mr. Feng said it very well, that the rates are very low in all places. But if you're philosophically want to focus on and say that, you know, it's three and half times, you know -- you can -- if the worst place is, you know, only performing at less than a third of the best place, you know, is that a significant gap. Maybe.

I mean, it depends how you look at it. I'll throw it out to the committee. How should we be looking at this, how have we looked at this historically?

Somebody's mentioning sentinel events, never events. Yeah, so you can -- naloxone administration's probably not a never event, I would say respectfully. You know, it's not defined by the NQF as a never event. It's not defined by the Joint Commission as a sentinel event. But I'll leave it to others.

Co-Chair Septimus: Okay, Matt, I don't know if you want to comment on the NQF perspective on this.

Mr. Pickering: Sure.

Co-Chair Septimus: That might help the committee a little bit.

Mr. Pickering: Well, when you're looking at performance gap as part of the opportunity, or the importance to measure a report, you're looking at this opportunity for improvement. So by just looking at the variation across different testing sites, or those being held accountable to the measurement as providers, accountable entities. It's also looking at the overall, or less than optimal performance.

Is there less than optimal performance in which you can argue whether or not these percentages are less than optimal performance for this measure to have a gap? And then there's also the disparities portion as well. Is this measure identifying disparities as various different subgroups, or subpopulations that there is a gap? So, again there's the disparity piece across subpopulations, the variation around accountable entities, or sites that are being tested, and then just looking at the overall less than optimal performance.

Is there just, the performance on this measure, there's room for improving overall. But I think I was interrupting someone when they were trying to chime in before I talked, so I wanted to make sure if someone from the standing committee is wanting to talk about any of this.

Co-Chair Septimus: Any hands?

Member Fairbanks: Well, I just want to comment on the naloxone administration, and the never event discussion. I think it is a never event to have somebody die from an opiate overdose on the floor, and the important thing is that naloxone is a save when everything goes right. I think we have to think of it as a proxy measure, a little bit, for something could go wrong, and we are the safety committee, so it's really important to think about it from what I would call a safety engineering perspective.

Member Yealy: Do you think it's a good proxy for safety?

Member Fairbanks: Do I think that naloxone administration is a good proxy for risk of safety of opiate overdose? Absolutely. Because if you have a place where you have a higher number of, and this is unscheduled, this is not counting during a sedation, you're titrating a little bit. This is when

there's been an emergency naloxone on a floor, I think it is a proxy for a risk.

Co-Chair Septimus: Jason had his hand up, Jason. I saw this one.

Member Falvey: Thank you. I've been following the discussion, and certainly can appreciate from the medical side, I'm going to comment from the statistical side as kind of a statistics nerd. But I'm kind of looking at this, I'm looking at the actual raw numbers from each of the sites. So what do you tell of site number two, that had seven events out of 2089 encounters, and they have a performance gap or rate of .34 percent?

You go to hospital site three, which had eight events out of 1784, and they're one standard deviation higher. They had one more event, and a few hundred less patients, I don't know, is that, you have eight events, what do you tell that hospital, and say you can do better?, we want to see that be six. I think it's really challenging when it's eight situations, and one or two patients throws that rate off by a standard deviation on the measure, what I'm looking at now, I think that is really challenging.

Member Charbonneau: But that one or two patient could have died, so there's that.

Member Pollak: So, if they don't administer naloxone, they meet the measure actually.

Co-Chair Septimus: Okay, let's do this, I think great discussion. Unless someone's got a brewing thing, we need to probably vote on whether or not the committee thinks there's enough of a gap to move forward. So, if I don't see any other hands, let's vote Isaac.

Mr. Sakyi: Now open for measure 3501E --

(Simultaneous speaking.)

Mr. Pickering: Wait, sorry, sorry. I was talking on mute, sorry Isaac. Don, did you want to share your comment at all? I just want to make sure if you wanted to share anything, really what you said in the chat.

Member Yealy: No, I think I'm flogging a dead horse.

Mr. Pickering: Okay, sorry Isaac, go ahead.

Mr. Sakyi: Voting is now open for measure 3501E on performance gap. The options are A for high, B moderate, C low, and D insufficient. Waiting for some more votes, one more. Voting is now closed for measure 3501E on performance gap. We have zero votes for high, seven for moderate, five votes for low, and four insufficient.

Co-Chair Septimus: Okay, I need some percentages here. It looks like we have nine low, and insufficient, versus seven. If my arithmetic is correct, that's less than 50 percent?

Mr. Sakyi: That's correct.

Co-Chair Septimus: Is it less than 40 percent?

Mr. Pickering: No.

Mr. Sakyi: 43 percent, consensus not reached.

Co-Chair Septimus: Okay, so that's a consensus tie, if I understand again, I'm pretty sure we're going to continue on, but this is going to be consensus not reached, correct?

Mr. Pickering: That's correct, consensus not reached.

Co-Chair Septimus: Okay, all right, let's go on then

to reliability.

Member Yu: Remember, I'm muted. The reliability and testing, basically for theta reliability, the developer compared electronically captured data to manually abstracted data from patient medical records. And they do the test using coefficient to qualify the (audio interference) the coefficient was 198 at one site, and one at all other sites for the six randomly selected subsamples.

The preliminary rating on the reliability is moderate, and the committee member's comments have, most of the members have no concerns about the reliability and testing. One member was concerned about the overflow agendas, and how bad would a change, the change would be supported in the hospital. Based on the test, I felt personally the reliability is pretty high. Any comments from my co-discussants?

Co-Chair Septimus: Ed, Sara? This one I think is less controversial. Any comments? Do you see anything Matt?

Mr. Pickering: No.

(Simultaneous speaking.)

Member Yealy: The only thing I would say at this time is that these reliability data are so good it gives me a little bit of pause. But we vote on what's in front of us, we don't speculate about it.

Co-Chair Septimus: It's a great kappa score.

Member Yealy: It passed great, and it's in the stupendous area.

Co-Chair Septimus: Okay.

Mr. Pickering: So, sorry Ed, I do see Jason has his

hand raised, and Jason, I'm not sure if that's from previous, or you have something to comment. Okay, I'll mention again, that this was reviewed by the SMP, it's an outcome measure, so it's reviewed by the Scientific Methods Panel, so it did pass the SMP. And what you'll see is the question around whether you want to uphold the SMP rating if we go to vote. But I don't see any hands raised, and no other questions in the chat box Ed.

Co-Chair Septimus: All right, well then let's --

(Simultaneous speaking.)

Member Yu: Could I --

Co-Chair Septimus: What's that?

Member Yu: I've got one question, could I ask a question?

Co-Chair Septimus: Sure.

Member Yu: I'm thinking the developer stated that about eight percent of facilities when tested for feasibility, they identified that that would be difficult to obtain data required as element. So, it mentioned work flow modification, so I'm just interested in the developer's thoughts on how this measure can be implemented consistently when you think about you have to change the overall work flow, and other difficulties.

Co-Chair Septimus: Is that, where would that best fit with --

Member Yu: This is under the reliability hub.

Mr. Pickering: So, this is actually part of the feasibility scorecard, which gets --

(Simultaneous speaking.)

Co-Chair Septimus: Yeah, I thought it would be feasibility, that's why I was asking.

Member Yu: Okay, then I'll hold off that one.

Co-Chair Septimus: It's all the feasibility.

Member Yu: Sorry about that.

Co-Chair Septimus: Okay, let's continue to vote.

Mr. Pickering: Sorry.

Co-Chair Septimus: Again, it's question number three, Isaac did you, I'm sorry I forgot, did you state the question?

Mr. Sakyi: Not yet. Voting is now open for measure 3501E, and the question is do you accept the Scientific Methods Panel's moderate rating for reliability? The options are A for yes, and B for no. We're waiting for two more votes. Voting is now closed for measure 3501E on accepting the SMP's moderate rating for reliability. We have 16 votes for yes, and zero for no. the standing committee votes to uphold the SMP's moderate rating for reliability.

Co-Chair Septimus: Thank you. We're going to move to, we heard a little bit about this earlier Yanling, and there's also a comment in the chat box that some of you may want to read. But why don't we go on to validity. Yanling? You're on mute I think.

Member Yu: Yeah, I just was on mute. The developer conducted two validity testings. One is the data element level validity, agreement between data electronically extracted from the sample patient EHR data, and data manually extracted from medical records was 100 percent, matching for our six test sites, except two. And the measured score validity using the status of positive predictive value

sensitivity, negative predictive value sensitivity to determine the percent of accuracy between the EMR data, and the abstracted data.

The PPV was 100 percent, and the sensitivity is 100 percent in all but one test site, and NPV is about 100 percent, and specificity is 100 percent. So, the accuracy actually was very high for both PPV, and NPV. So, the score level, and the EHR reported opioid related adverse events was compared to clinical review of the patient records. As far as the missing data to the validity, the developer does not anticipate any missing data, or basically they think it's nearly zero.

The preliminary rating is moderate, and the SMP rating is pretty mixed, it went from low to high. And the committee member's comments, most members expressed no concerns for validity testing, and threat of validity, and think no risk adjustment is appropriate. But one member expressed concerns over a small number of a test site, and one member thinks a risk adjustment may be needed, as certain conditions may increase risk.

The same member also questioned why no exclusion for patients on hospice, cancer, sickle cell, et cetera, and whether using a rate per hundred patients to protect hospitals that use opiates appropriately, but the results are skewed by a small number of events. So, this is what we have for validity.

Member Sood: May I --

Co-Chair Septimus: Patrick had mentioned some of this earlier in his comments. Okay, who wanted to speak? I'm sorry, I didn't see. Oh, Geeta.

Member Sood: Sorry, I just wanted to ask a clarifying question, which is also a hidden comment

I guess. When we talk about validity, I know the scientific methods panel focuses a lot on statistical validity. We also mean clinical validity, correct?

Mr. Pickering: Yes, that's correct Geeta.

Member Sood: Okay, thank you.

Co-Chair Septimus: Okay, any other comments? Again, some of this we've heard. Don, I thought you were going to raise the issue 100 percent positive predictive value, 100 percent.

Member Pollak: So, I would just repeat what, I can give Don a chance, I'm sorry, go ahead.

Member Yealy: No, go ahead, you're good.

Member Pollak: The issue I was going to raise is, the way I read it, and I may have misread it, and I know there are people from the developer among my colleagues, but it looked to me like what it was validating was through the EMR, they say that naloxone was given after an opioid, that's in fact what happened. I think what Geeta and I are saying is that something bad also happened, like the patient had respiratory depression.

Then that's where we get into the discussion that we had earlier with Dr. Romano, where I was saying I've never in my life seen the ability to have such good concordance, even with STS abstractors who are trained. I don't know where you trained these people that they get such good ability to tell that naloxone is being used appropriately. So, good for them, but I don't know that, where that independent meter comes from. Where is the standard meter here? I'm struggling with that.

Ms. West: Just to quickly address a little bit of that, in the testing, they looked at a number of things. They looked at the nursing notes where they

documented what was going on with the patient at the time of the administration. They looked at the POSS scores that look at their, basically kind of their standard documentation that they do to check on the status of patients that have been administered opioids.

We looked at both the scores that occurred immediately prior to the event, the scores that were documented afterwards, and then also looking at the free text notes in order to ascertain that information.

Member Pollak: Is there not a chance of bias in light of naloxone having been administered, and people thinking therefore that it was respiratory depression?

Ms. West: We can't speak for the sentiment of the nurses that were documenting it, and with some of these instances we had our staff that were actually doing that documentation, and then in other instances they were independent evaluators that were outside of the development team themselves. So, we didn't have that level of bias, it was one person, or one entity that did the calculation from the electronic, and then another entity that actually did that extraction of the data.

Member Pollak: Thank you.

(Simultaneous speaking.)

Dr. Feng: This is Bo from IMPAQ International on the developer side.

Ms. West: Go ahead.

Dr. Feng: We definitely share your concern, and first of all thank you very much for the comments. To be honest with you, if AI is so advanced, natural language processing is so advanced, then these

(audio interference) can definitely incorporate that kind of nursing nodes, natural languages into the calculation, but we are not there yet. But we definitely tested empirically the extent to which false positives can really affect, or impact the measure.

The extent to which the measure is currently specified, along with time windows of data value sets can truly detect a true positive, and we did that. And through nursing nodes, we found a majority of cases, the indication for the Narcan, as well as the patient reaction after receiving the Narcan, and we found that with a majority of the cases Narcan was used to reverse the opioid, or excessive use.

Co-Chair Septimus: Okay, so we're on validity, does anybody else have any comments on validity? Again, this is an E measure, just to remind everybody.

Mr. Millet: Hi, this is Chris from NQF staff, can you hear me?

Co-Chair Septimus: Yeah Chris.

Mr. Millet: Yeah, just one other point to bring up on validity, and this issue is identified in the feasibility scorecard, so it's probably easy to miss. But in the scorecard the developer indicated that one of the data elements had an issue with accuracy, and just wanted to circle back here to bring up if there's any impact, how does that relate to the validity of that data element? This might have been an issue again, that you were speaking to earlier.

Mr. Pickering: Yeah, this is that encounter performed, encounter inpatient facility location operating in room suite data element.

Ms. West: Yeah, so in terms of that particular data

element, in terms of location, in order to capture whether or not the patient was in the operating room, whether that's for the exclusion part, or however to determine that they were in the operating room when they received the Narcan to exclude them from the measure, since it's not capturing that. Some facilities, the locations aren't actually tracked, or not tracked easily.

So, the sites that we looked at, we looked at 23 sites, only 2 of them had an issue with being able to track those temporary locations. So, that particular item, we didn't feel that was a significant enough issue. Some of them were able to use proxy information, for example the location of the administering staff in order to determine where that patient was in there. So, we don't feel that that particular element is a significant issue in order to be able to flag these patients.

Co-Chair Septimus: Chris, does that answer your question?

Mr. Millet: Yes, it's more just to have the developer explain that, and see if that has any reaction from the committee.

Co-Chair Septimus: Any other comments?

Member Fairbanks: Well, I feel the need to kind of close the loop on the earlier discussion I was involved in, in view of this discussion. This discussion does concern me a lot, about what people have raised about whether this could inhibit the use of Narcan in a time when it's diagnostic, or when it may be necessary, but not clearly necessary, and I think that does, I think it's relevant to the validity discussion.

Co-Chair Septimus: So, Terry you believe, this horrifies me by the way, if true, you believe that to

improve your standing, that you would not administer potentially lifesaving medication, so you would look better in terms of reporting?

Member Fairbanks: No, this is what, Edward said something that, what you just said is not my point. I think that the issue is not when you have somebody apneic in front of you that needs Narcan. The issue is when Narcan is used diagnostically, and it sometimes is helpful to patients. And second, I don't think an individual provider is going to make the kind of decision that you're talking about.

I'm concerned about a cultural transition because of misunderstandings of leadership that then prohibits an individual clinician, or makes it harder for them to do the right thing. That's more what I'm concerned about over time.

Co-Chair Septimus: No, no I'm understanding, I'm just saying that that concept horrifies me actually. Ed, you were going to make a comment?

Member Pollak: Well, very respectfully, people look at these measures who are hospital administrators, and they're not clinical, and they set protocols up, and they assume that there's really solid evidence that it's a bad thing, that what these measures are, that doing better on these measures matter. And I completely agree with, Terry put it perfectly. No one is saying that somebody apneic in front of them, and somebody says should we push Narcan? And they say no, we're going to fail the measure, that will never happen, and if it does, somebody mentioned malpractice, agreed.

The issue is similar with issues we get into of blood cultures to avoid, and we all know this happens, to avoid certain hospital acquired infections. It doesn't mean the measure on the whole is bad, this isn't an all or nothing about this measure, it just is reality,

which is that people need to be judicious about it, and have appropriate exclusions, and so on. So, I don't think that anybody is saying that our practitioners are evil, and only care about meeting these measures.

But protocols matter, and there's not solid evidence that it's worse. There's not good evidence that it's better to do what Don and I were suggesting, which is when it's unclear, give Narcan, but that's the current practice pattern, and this I think gives hospital administrators a reason to not follow that practice pattern, which I'm not --

(Simultaneous speaking.)

Mr. Pickering: So, sorry Ed, and everyone, I just want to make sure that we keep our discussion here on validity, but.

Co-Chair Septimus: I was overstepping, we still need to, I got this Matt. Let's vote on, you're raising great issues, and it's going to come up at the end whether or not you think this is a suitable measure, but the question is the validity. That's what we're trying to vote on now. So, I would suggest that we vote on validity, and finish this so we can, because we're going around in circles, going almost all the way back to the evidence, and unintended consequences, we need to discuss about validity. So, let's vote on validity.

Mr. Sakyi: Voting is now open for measure 3501E on accepting the Scientific Methods Panel's moderate rating for validity. The options are A for yes, and B for no. We have 16 votes, voting is now closed for measure 3501E on accepting the SMP's moderate rating for validity. We have 63 percent voted yes, and 38 percent voted no.

Co-Chair Septimus: Okay.

Mr. Sakyi: The standing committee votes to uphold the SMP's moderate rating for validity.

Co-Chair Septimus: Feasibility.

Mr. Pickering: So, and actually this is why I was trying to chime in, I think we actually need to stop with the measure today, and continue on with feasibility use, and usability for tomorrow. The reason being is we have a number of people that have been waiting to provide public comments today, and so we do have time tomorrow to pick this back up. So, I do apologize to you, the developer, IMPAQ, but we will be reconvening tomorrow, and we can pick back up with 3501E on feasibility. So, Ed, if that's okay, we can --

Co-Chair Septimus: Yeah, two choices, one, can we get through the rest of this, and then have people stay over a little bit longer, but I think it's probably going to require more discussion just based on what's happened in the last 30 minutes. So I think, I don't want to rush this discussion, it's too important, there's still some issues we need to noodle on, we can do that overnight. But I agree, if it's okay with the developers, that we go on to the public comment, and then pick it up tomorrow at feasibility, okay?

Does that make sense, is everyone okay with that? I know this has not been -- I'm a little frustrated, I wish we could tie this discussion a little bit better, I think that the comments are excellent, and we'll get back to it again tomorrow. Okay, public comment then.

Mr. Pickering: Great. Thank you all very much for your time. We do have time tomorrow, and that's why we anticipated we would probably need to use it. So that's 2 p.m. Eastern tomorrow when we'll pick back up, 2 p.m. Eastern tomorrow. But now we

have an opportunity for the public to make any comments related to the previous measures that we have been discussing today.

So, if the public would like to make a comment, if you'd like to just chime in, and say your name, and where you're from, and please make your comment. And there's a couple hands raised, I see Tom Heymann, is that right? Would you like to make a comment?

Mr. Heymann: That's right, yeah. I'm trying to be polite by raising my hand, not just jumping in.

Mr. Pickering: Thank you.

Mr. Heymann: Thank you, I know you guys have had a very long day, and we started it with you. My name is Tom Heymann, I'm the president, and CEO of Sepsis Alliance. I want to thank you for your re-endorsement today of the SEP-1 measure, and just speak with you very briefly about what this re-endorsement means to the Sepsis Alliance community. Sepsis Alliance was founded in 2007, we are the nation's leading sepsis patient advocacy and education organization.

We represent sepsis caregivers, survivors, patient advocates, and the families of those who have been lost to sepsis. We serve more than 2.5 million people each year. The organization was founded by Dr. Carl Flatley, who tragically lost his healthy 23 year old daughter Erin to sepsis after a routine procedure in the care of their local hospital in 2002. That was long before the SEP-1 measure was first endorsed by this committee.

It was long before hospital leadership was focused on the early detection of sepsis in the way that we are today. There are still too many urns out there, and too many of their deaths are preventable.

Sepsis is common, deadly, and devastating. It causes 14,000 new amputations every year, and thousands, and thousands of cases of post-sepsis syndrome. In fact one in five people diagnosed with sepsis will die in hospital, or be transferred to hospice.

It's personal, one in three of us will get sepsis in our lifetime. 1 in 15 of us will die of it, likely before the time we should have passed. In underserved communities, and communities of color, the burden of sepsis is even greater. Black Americans bear nearly twice the burden of sepsis deaths compared to White Americans, and sepsis moves fast. Each hour of delay before a septic patient is treated is associated with a four to nine percent increased risk of mortality.

Early diagnosis, which SEP-1 encourages, is the key to saving lives and limbs. The Sepsis Alliance community notes, very often first hand, that effective sepsis care is all about time, that's why it matters that this committee has re-endorsed this measure today. And we agree with what today's lively discussion revealed, that SEP-1 must continue to evolve as we refine sepsis screening, develop new diagnostics, and devise optimal treatment methods.

We still have a lot to learn about sepsis. We also acknowledge the importance of curbing antimicrobial resistance, but we can embrace these two responsibilities together, treating sepsis patients in a timely way, and combating AMR. SEP-1 focuses on the first dose, and then encourages responsible stewardship. The Sepsis Alliance community agrees, 1,000 of them signed a petition just this week in support of the re-endorsement of SEP-1, which you delivered on today.

There is strong unified sentiment from sepsis care

givers, survivors, and those closest to them to keeping this measure in place, to keep our collective focus on sepsis was the right thing to do, and we thank you for your vote today. But we need to remember that the care you all guide is about real people, and with your permission, I would like to now give the virtual microphone to Katy Grainger, who is a sepsis survivor, and a Sepsis Alliance board member. Thank you so much.

Co-Chair Septimus: Tom thank you, as always very eloquent, and passionate in your cause.

Mr. Heymann: Thanks Ed.

Ms. Grainger: Thank you so much Tom, can you guys hear me okay?

Co-Chair Septimus: Yes.

Ms. Grainger: Hi, I'm Katy Grainger, and I am a sepsis survivor. Two and a half years ago I went into septic shock, and I became a multiple amputee as a result of that because of a small infected cut on my finger. I'm now a sepsis patient advocate, and community educator. I'm here today to thank all of you for your support of SEP-1, you made my morning. It's so great to see that measure pass again, because I believe it saved my life.

When I first noticed the unusual infection it was purple, and kind of oozing, that was on my finger, that led to my illness, I had no idea how to recognize the signs, and symptoms of sepsis. Because the infection was so strange, I had it examined at a walk in medical clinic near my home, which was on the remote island of Kauai in Hawaii. My vital signs were still strong at that time, so I was given a topical antibiotic, and I was told to pick up a prescription for oral antibiotics that I should begin taking if my infection would get worse over the

weekend.

As you know, sepsis is extremely difficult for hospitals to diagnose, and even harder for patients to recognize as they are becoming ill. Because of this, I arrived at our small community hospital 36 hours after I had visited that walk in clinic, having told my friend who was driving me, that I had never been so sick. My blood pressure was 50 over 30, and I was complaining that my hands and feet were burning.

But the medical staff was unable to diagnose me with sepsis at that time. Because my symptoms were indicating the possibility of sepsis, my clinicians followed the protocol that was required by SEP-1. Among other things, this meant that I was given fluids to try to elevate my blood pressure, and I was immediately started on broad spectrum antibiotics to fight the infection that was causing me to become so ill. It would be several hours before it was determined that I was indeed in septic shock.

Two days later, after being air lifted to a level 1 trauma center in Honolulu, my husband was told that I was receiving the best medical care possible, because I was literally the sickest person in the entire hospital. Despite their best efforts, he was told to call my children, as I might not survive. And they were all prepared that if I did survive, I would likely lose my hands and feet.

When I awoke a week later, every doctor that visited me in my room over the next couple days had a huge smile on their face telling me that I was a miracle, and that they were so happy that I had survived because I had beat the odds. Had my medical team on Kauai delayed my sepsis treatment at all, I would likely not be here to thank you all for your support of SEP-1. I didn't know anything about SEP-1, but I am grateful that my tiny community

hospital did.

The diagnosis, and treatment that saved my life were in no small part attributable to this measure. As multiple members have brought up today, SEP-1 kept my doctor's attention focused on the possibility of sepsis, and it encouraged their early intervention. Without SEP-1 I would likely have been on the wrong side of the devastating sepsis survival statistics.

As it is today, I've lost both my lower legs, and several of my fingertips, but I still have my hands, and I am alive. Thank you again for your support of SEP-1 so that more lives can be saved from this deadly condition.

Co-Chair Septimus: Ms. Grainger, thank you so much for sharing that. Very, very gratifying to know that you survived, and hopefully we can do better, so thank you for your comments.

Ms. Grainger: Thank you.

Mr. Pickering: I also echo what Ed has shared, thank you so much as well for sharing with the committee. And I'm just going to make a correction on myself here, sorry, it's been a long day. If you are from the public, and you'd like to chime in to speak in front of the committee, please raise your hand. There's a raise hand feature, if you're not able to locate it, identifying where your name or phone number is within the chat platform, and then you can hover over, there's a raise hand option there, you can raise your hand.

So, I'm just going to go down the list here. Rebecca Smith-Bindman I believe, is that you?

Dr. Smith-Bindman: That's right. Thank you very much, thank you for the opportunity to speak. I am a radiologist at University of California San

Francisco, and I'm commenting on the ACR measure. The radiation doses used for CT examinations are highly variable across hospitals, and imaging facilities. For a patient's image for the same indication, they're frequently far higher than needed for diagnosis, and they're in the range we know to cause cancer.

A quality measure that can inform clinicians about how they can safely lower radiation doses used for CT scanning, while maintaining image quality can greatly improve the health, and safety of patients. However the ACR's proposed measure is inadequate for this purpose, and if it goes on to adoption, and use, could undermine the broad application of much more effective ways of using quality measures to achieve this goal.

The primary concern that I have is the validity of the measure. The most important predictor of radiation dose is the choice that radiologists make as to what protocol should be used. Should a patient be scanned with a single phase, or a double phase scan? Protocols with more phases deliver proportionally more radiation, yet in most indications there's no evidence suggesting that higher phases produce better diagnostic utility.

The ACR defines the target population for the measure as all patients who require either a CT abdomen and pelvis single phase, a chest exam with single phase, or a CT head brain with single phase. But the measure does not identify patients who require these exams based on the clinical need, but were in fact given much higher doses through multiple phase exams when a single phase study would have been the proper decision.

This limited assessment of dose within these stratified groups ignores the primary factor determining the patient's dose, which is which

protocol to use, which is almost entirely at the discretion of the imaging physician. The measure will only assess the relatively smaller variation in technical parameters within these protocols, but will leave unassessed the variation that occurs to the choice of protocols.

The unnecessary variation in protocol selection is the critical factor, but the ACR measure over adjusts for this by stratifying this measure based on the protocol. In the University of California San Francisco CT dose registry, which includes 8 million CT scans collected from 162 hospitals and imaging facilities, these three exam types overall account for over 39 percent of CTs, so a meaningful number.

However, these three exam types account for one percent to 83 percent of CTs across the different hospitals and imaging facilities, suggesting the denominator for this measure is variable, and does not reflect a patient population who require these exams, but rather reflects the varying decisions of the radiologists to assign patients to different protocols. The only way to accurately judge physicians, and their use of radiation for CT is to evaluate how they use radiation in the population of patients where the selection of imaging protocol entirely under their control is included in the assessment.

Radiation doses must be assessed by examining the intent, and clinical question of the provider ordering the scan, not based on what the radiologist chooses to do, which is subjective. The measurement of the dose within these narrowly defined groups as defined by the ACR will only serve to camouflage the large variation in practice that exists, 100 fold variation, and will not serve to improve practice. Thank you for the chance to comment on this measure.

Co-Chair Septimus: Thank you so much. Matt, of course we do have a post comment discussion, and I think Matt, either we record that, or should that be put in writing in the post comment, so we can discuss that at the post comment session? What would be the best way for her to address that?

Mr. Pickering: Right, so Rebecca, if you'd like to submit your comment as well during the post comment period, those comments will also then be considered as well by the standing committee during post comment when the measure goes to post comment.

Co-Chair Septimus: Is that okay?

Mr. Pickering: There is an opportunity to do so, yeah.

Co-Chair Septimus: We'd like to have that in writing, so we can discuss that in the post comment, is that okay?

Dr. Smith-Bindman: Of course, thank you.

Co-Chair Septimus: I appreciate your thoughtful comments. Okay, who's the next one?

Mr. Pickering: So next on the list with their hand up, is just Steven Simpson, or Dr. Simpson.

Dr. Simpson: Yeah, thank you. Yeah, my name is Steven Simpson, I'm a professor of medicine at the University of Kansas in Kansas City. I've spent a large part of my 30 year career in medicine educating smaller hospitals about how to recognize sepsis, and how to treat sepsis early, and aggressively, and how to do quality improvement, so that they can demonstrate that they're doing exactly that. I know that you are all aware, because you all reviewed studies from such places as Johns Hopkins, and Harvard, and the University of

Pittsburgh Medical Center showing that SEP-1 measure did not contribute to their overall survival rates in sepsis.

But I can assure you that it has improved the overall mortality, and the quality of care in the sorts of hospitals that I've been educating throughout my career. I've been to several hundred, not several hundred, I've been to over 100 hospitals around the nation, many of them in Kansas, where I live, helping to educate, and I can say that before 2015, not a single hospital that I ever visited had any sort of structured treatment, and diagnosis of sepsis in place.

So, I just wanted to say thank you for continuing this measure, it's not the UPMCs, or the Johns Hopkins that necessarily need it, but it's all of those guys, and they have benefitted tremendously, so thanks.

Co-Chair Septimus: Dr. Simpson, thank you so much for hanging in there to provide that comment, and waiting.

Dr. Simpson: Yes.

Co-Chair Septimus: Matt, who's next?

Mr. Pickering: So, I see a hand, all the name is Dana, or Donna I believe, sorry if I'm mispronouncing your name, it's the hand that's raised.

Ms. Wollins: No problem, yes, and this is Dana Wollins, I hope you can see me now, and I wasn't able to get my last name up on the reader. I'm from the Infectious Diseases Society of America, the IDSA. I'm vice president of clinical affairs, and guidelines, and I wanted to speak on behalf of the organization today. First off I want to thank the NQF for this opportunity to provide comment. I'm

wishing to speak on NQF measure number 500, the SEP-1 measure bundle.

I want to share some remarks that are on behalf of the IDSA sepsis task force that includes, as its partner organizations, ACEP, the American College of Emergency Physicians, the Pediatric Infectious Diseases Society, the Society for Healthcare Epidemiology of America, the Society of Hospital Medicine, and the Society of Infectious Diseases Pharmacists. IDSA was disappointed not to have an opportunity to address NQF before the vote, and written comments, we appeared not to have the communication quite right in being in the loop on the deadline for submissions, hence we want to make sure to file some remarks now, and we'll make sure if there's post communication we're able to do in writing we will as well.

IDSA acknowledges the success of SEP-1 in raising awareness of sepsis, and in focusing measures that could improve patient outcomes, and we fully support the importance of measuring sepsis, and a performance measure on this critical issue in healthcare. However this NQF review presents an appropriate time to issue IDSA's significant concerns regarding this measure. Namely that there is a lack of consensus that adherence to the bundle leads to better outcomes for patients, that the measure carries unintended consequences, including overuse of antimicrobial agents, and antibiotic adverse effects.

And that it brings with it an out sized reporting burden on hospitals and providers. IDSA has published a position paper with the aforementioned organizations in 2020 that outlines concrete concerns, and recommendations to improve this quality performance measure. We hope the real concerns that would be noted by NQF, and CMS, and we recognize that measure is far more powerful

if providers have the most confidence in the measure being placed into practice.

We would like to see separating septic shock from sepsis without shock. We would like to see redefining of time zero. We would like to see making this bundle an E measure, and we would like to see assessing new, and emerging data on the impact of the bundle on patient outcome. We're pleased to be collaborating productively with the SEP measure 1 stewards to evaluate alternative definitions of time zero, and timing of antibiotic administration currently.

We believe this is a good, constructive beginning that IDSA hopes to continue to work with the stewards, and CMS to improve the bundle, and the confidence in stakeholders that the invested efforts truly translate to improve outcomes. Thank you very much.

Co-Chair Septimus: Thank you Dana, and as the previous speaker, it would be again, helpful, if IDSA would comment in writing during the comment period, so we can consider those comments at our post comment meeting, thank you very much.

Ms. Wollins: Thank you very much Dr. Septimus.

Co-Chair Septimus: Any other hands Matt?

Mr. Pickering: No, I don't see any other hands besides the ones that have been raised already.

Co-Chair Septimus: This may be a record, this is great. We have public comments, and generally it's silent. So, the good thing is that our patient safety community is alive, and active, and interactive, because I think out of interaction, and constructive respectful conversation, we make things better, and I think I heard that over, and over again during the day today. It's been a long day, but a very

productive day. We'll take up again tomorrow at, is that 1 p.m. Eastern Time?

Mr. Pickering: Yeah, so I'll just interject. Thank you as well for all of your comments. I agree with Ed, this has been great to hear the public voice any concerns, or (audio interference) for the proceedings today, so thank you very much. As Ed was mentioning, we do have a meeting tomorrow, it starts at 2 p.m. Eastern, it goes until 5 p.m. Eastern, and we will pick up where we left off with 3501 starting with feasibility, and then we will finish up with 3389, which is a PQA measure.

And then there's still the related, and competing discussion. So, we'll pick up tomorrow from where we left off today. I want to thank everyone for their time today, it's been a long day. Standing committee, the developers, those members of the public, thank you all so much for your time, and also thank you to the NQF staff as well for all of their effort, and giving us support too today.

We're not over yet, we will reconvene tomorrow, and get some rest, and thank you all very much. If you have any questions, please don't hesitate to reach out to us at the project box, thanks everyone.

Co-Chair Septimus: This is the best committee at NQF Matt, is it not?

Mr. Pickering: That's it, no bias there Ed, no bias there.

Adjourn

Co-Chair Septimus: This has always been the best committee, I'm sorry Owen isn't here, but we have shared so many moments, and learned so much from our committee members over the, literally almost close to 10 years that we've been involved with patient safety, and it's just amazing with new

blood, and new people, how much better it gets. So, it only gets better because of your engagement, and volunteering because sometimes the work is hard, but I think the results are rewarding. So, thank you all very much for your time, and I'm sure I speak for Iona, she would send her best as well.

We'll reconvene at 2 p.m. Eastern Time, 1 p.m. Central Time, because I'm on Central Time. I know that Yanling is on Pacific Time. So, have a great afternoon, folks. Bye-bye.

(Whereupon, the above-entitled matter went off the record at 5:10 p.m.)