NATIONAL QUALITY FORUM

Moderator: Benita Kornegay Henry
June 19, 2019
10:36 am CT

Desmirra Quinnonez: Good morning everyone and welcome to the Behavioral Health Substance (Use) Measure Evaluation meeting. This is the first of two of our web meetings, and we are happy to have you all participating on the call. So we'll go ahead and get started. Just as a reminder, if you could please mute your lines if you're not using - if you're not speaking. That will help to eliminate a lot of background noise of the call.

Just want to introduce our project staff. So we have Michael Abrams here, the Senior Director. Nicolette Mehas, who is on the line as well, is here in person with us. I'm Desmirra Quinnonez. We have - and we actually have a new member of our team. Hopefully she'll be permanent. This is Hannah and she's been working with the behavioral health team over the last couple of weeks. And we also have Elisa Munthali who is our Senior Vice President on as well.

So just a brief overview of today's agenda. I know today is a little unconventional and the way we've done our meetings in the past, but we'll have two web meetings today. This is the first of two from 11:00 to 2:00. We

try to give you a brief break in between so that you can grab some lunch and stretch your legs, and then we'll reconvene at 2 o'clock.

So for this first meeting, we're going to be discussing the measures that you have before you and doing a brief overview of our evaluation process. And for the next meeting, we will continue. And hopefully if we have time, we will be able to discuss a few portfolio gaps and have our related and competing measure discussions.

So at this time, I will turn it over to Elisa so she can go over our introductions and disclosures of interest.

Elisa Munthali: Thank you, Desi, and good morning everyone, and thank you so much for joining us this morning for the two web meetings. What I'll be doing is - I think there was some feedback. Sorry about that. I will be combining introductions with disclosures of interest.

And just as a reminder, when you were named to this committee, we asked you a number of questions that they're related to work that you may have done outside of NQF, that's related to the work of the Behavioral Health Committee. And so what we're asking you to do today is to orally disclose any information that you've provided to us on the forms that we sent to you, but only as it's related to the work in front of you.

Just a couple of reminders. You sit on this committee as an individual. You do not represent the interests of anyone who may have nominated you for the committee, or your employer. We are interested in not just those activities that are relevant to the work that are paid, but also those that are unpaid.

And perhaps the most important reminder is, just because you've disclosed,

does not mean you have a conflict of interest. We go through this process in the interest of openness and transparency. And so what I will do is I have the roster, starting with your co-chairs, and then listed in alphabetical order as it is on the screen.

So I will start with Peter, ask you if you introduce yourself, tell us who you're with, and if you have anything to add.

Peter Briss: Hi. This is Peter. Welcome everybody and - or welcome back I suppose. I'm still Peter Briss and I have nothing to disclose.

Elisa Munthali: Thank you, Peter. Harold?

Man 1: He's not on.

Desmirra Quinnonez: He had to jump off for a graduation but he'll be coming back on a little bit later.

Elisa Munthali: Great. Then just as a note here, Harold is conflicted on - and my eyesight is really bad, measures 3388 and 3399. These are the NCQA measures. Okay. Jeff Susman.

Jeff Susman: I don't have anything to disclose. I'm still doing the same things that I was doing.

Elisa Munthali: Thanks, Jeff. Mady. Mady, are you with us? Are you on mute? Okay. David.

David Einzig: Yes, I'm here. Good morning. No disclosures.

Elisa Munthali: Thank you. Julie.

Julie: Hi. How are you? This is Julie (unintelligible) and I have no disclosures.

Elisa Munthali: Thank you. Connie.

Constance Horgan: Good morning everyone. I am conflicted on the NCQA measures on behavioral health, the last two measures that are listed that will be discussed this afternoon.

Elisa Munthali: Thank you for letting us know. Lisa Jensen.

Lisa Jensen: I'm here. Nothing to disclose.

Elisa Munthali: Thank you. Dolores Kelleher.

Dolores Kelleher: Hi. This is Dodi Kelleher. I have nothing to disclose.

Elisa Munthali: Thank you. Kraig.

Kraig Knudsen: Hello. This is Kraig Knudsen and I have nothing to disclose.

Elisa Munthali: Thanks, Kraig. Charles Gross. Okay. Michael.

Michael Abrams: Hi. Morning everybody. I have nothing to disclose.

Elisa Munthali: Thank you. Tami.

Tami Park: Hi. Good morning. I have nothing to disclose.

Elisa Munthali: Thanks, Tami. Brooke Parish.

Brooke Parish: Yes. This is Brooke Parish. Good morning. I work for Joint Commission. However, I've never worked on any of the measures submitted by Joint Commission.

Elisa Munthali: Okay. Thank you for letting us know. David Pating. Okay. Vanita. Okay. Lisa Shea.

Lisa Shea: Yes. No disclosures.

Elisa Munthali: Thank you so much. Andrew Sperling.

Andrew Sperling: No disclosures. I'm here.

Elisa Munthali: Thank you. Michael Trangle.

Michael Trangle: Trangle, but I'm here and no disclosures on the stuff we're talking about today.

Elisa Munthali: Thank you so much. Bonnie Zima?

Bonnie Zima: Good morning. Nothing to disclose.

Elisa Munthali: And Leslie.

Leslie Zun: Good morning. I have nothing to disclose. I do have a new job, but there's no (unintelligible).

Elisa Munthali: Okay. If you could just let know who you're affiliated with, you can send us a message to the behavioral team, that'd be great. And …

Leslie Zun: Sure. Let me know what department.

Elisa Munthali: Okay. Thank you so much, Leslie. I think someone was trying to chime in?

Mady Chalk: Yes. This is Mady Chalk. Nothing to disclose.

Elisa Munthali: Hey Mady.

Mady Chalk: Nothing to disclose.

Elisa Munthali: Okay. Thank you, Mady. Before I turn over the meeting to my colleagues, I just wanted to remind you, at any time if you remember that you have a conflict, we want you to speak up. You can do so in real time via the chat function, or you can speak up by your speaker. Likewise, if you believe that anyone on the committee is acting in a biased manner, we want you to speak up. So I'll turn it over to Desi. Thank you.

Desmirra Quinnonez: Thank you, Elisa.

Vanita Pindolia: I just wanted to say, this is Vanita. I'm on the call as well.

Elisa Munthali: Thank you, Vanita. Do you have anything to disclose?

Vanita Pindolia: No. Nothing to disclose.

Elisa Munthali: Okay. Thanks.

Desmirra Quinnonez: Thank you. All right. So at this time, we'll turn it over to Michael who will give us a brief overview of the scientific method panel review.

Michael Abrams: Okay. So one measure was - so Michael Abrams here. Welcome everybody. Nice to virtually reconnect with you here as summer begins. So there was one measure that would have come to us had it passed methods panel review, but it did not. It's measure 3492, title emergency department use due to opioid overdoses, an obvious important subject at this time and our nation's history of course.

And just a reminder to all of you that NQF is now engaged in a technical expert panel process that is looking specifically at measures, many of which you all have reviewed and endorsed in the past that are germane to addressing that current crisis that is being faced in the US and internationally, but with focus of course on the US.

So this particular measure, although there was interest in it because of the topic, did not pass scientific methods panel review. And the principal reason by my memory was that it did not have ICD-10 testing specifications proffered. So somewhat of a small technical issue. There were other reasons as well, which if you're interested, we - can be discussed later. But that was the principal reason that it didn't pass.

We're hopeful that it will be - come up for a future cycle. So just wanted you to be aware of that. There were no other measures that went through the scientific methods panel relevant specifically to our discussion today or to our CDP in particular.

So with that, I'll hand it back to Desi to give you a quick reminder overview of

the evaluation process before we start to look at the measures that we're going to talk about today, this morning and this afternoon.

Desmirra Quinnonez: Thank you, Michael.

Tami Park: Hey Michael, can we - this is Tami. I mean, I'm very interested in that (unintelligible).

Michael Abrams: Yes. So I got - yes.

Tami Park: And I don't - can you kind of remind us how the review relates to our review? Like we're looking at liability (unintelligible).

Michael Abrams: Yes, of course. Yes, thanks for asking …

Tami Park: And what - yes.

Michael Abrams: Thank you for asking, Tami. And I did get your email this morning, so thanks for giving me the heads up. So the methods panel explicitly looks at one of the five criteria that you all consider. And there are statistics - that group is composed of statistical people with statistical and methodologic expertise in particular.

So they're looking at scientific acceptability specifically, which means of course reliability and validity. And if they feel that it doesn't pass, then the measure is held so that you all don't have to take time to review it until it gets prepared for future submission. That's the whole idea of the methods panel, to be a resource to you all.

Even if the measure passes, by the way, information does come to you from

the methods panel about what concerns and strengths and weaknesses that they observed as well to help you with your review. So does that clarify for you?

Tami Park: I mean not really because …

Michael Abrams: What I - any other …

Tami Park: I mean, the support - it feels like it's sort of undermining our judgement around what's reliable or not. I mean - or I would like to be able to understand more why they said that this, as you noted, critical measure is not reliable in those.

Michael Abrams: Yes. Well, so they're looking at specific quantitative issues and specific methodologic issues that presumably are cross-cutting, not specific to behavioral health or any other CDP that they'd be looking at. That's the idea. And, you know, as I said, a key reason that this measure didn't get through is because there was no ICD-10 testing, even if they were proffering to measure as ICD-10 ready and compliant, and the specifications included ICD-10.

And even as QNF has had in place a policy that triggered ICD-10 requirements for this current cycle and it wasn't met. So that was one of the …

((Crosstalk))

Tami Park: I mean I'm - so if that - that seems like a minor thing. And then - so could they just come back quickly update it to ICD-10 and then bring it forward or now (unintelligible).

Michael Abrams: Yes. So they …

((Crosstalk))

Peter Briss: So Michael, this is Peter. This is a really important issue I'd like us to table it so that we can get through the rest of today's agenda. Tami, if we need - I think this is an important issue for us to talk about. I'd rather not do it right now.

So let's - Michael, talk about - let's talk about ways to have this conversation to everybody's satisfaction elsewhere and outside of today and let's try to move on with the agenda please.

Michael Abrams: Yes. That sounds good. And then just as a final point for now, as it - again, so we have two submission cycles every year. So the idea here is that the developer would - could fairly quickly revise it and then submit it for the next cycle. So that's what we're hoping will happen in this case, and in cases …

Tami Park: Yes. I mean if you can send out notes from that meeting so we can understand. I mean (unintelligible).

Michael Abrams: Sure. I'll get more detail about that. Absolutely. I'll get that out to you. Happy to do that. Okay, but that's …

Tami Park: Yes. That would be good.

Michael Abrams: Okay. Great.

Man 1: I'd like to see the notes too.

Michael Abrams: You bet.

Desmirra Quinnonez: Okay. Thank you, Michael. So I know that you all are a very seasoned committee. So we're not going to take a lot of time for this overall overview of the evaluation process, but did want to give you just a quick reminder. I have some brief ground rules for today's meeting.

So we hope that you are prepared and have reviewed the measures beforehand so that you're ready for the discussion. And we'd like to base the evaluation and recommendations on the measure evaluation criterion guidance, which we included in the email that we sent out.

We would love for you to remain engaged without distractions. So if you could please send us a quick chat if you have to step away. And you can mute your lines if you're not speaking on the call. That would be really helpful. We want to remind you to just keep your comments concise and focused so that we can remain on topic and master the criteria that we're actually discussing at the time to keep us on task for our agenda.

And if you would announce your name prior to speaking, that would be really helpful as we're not in person and we can't use our visual cues. And if we could just avoid dominating the discussion and allow all others to contribute, that would be great.

So the rules for the standing committee during the meeting. Of course, you are to act as a proxy for NQF multi-stakeholder membership, and you've been working with NQF staff to achieve the goals of the project. So thank you for that.

And we're going to evaluate each of the measure against the - each criterion.

And so we will make - you will make recommendations regarding endorsement of NQF membership. And another goal of yours is to help oversee the portfolio, the behavioral health and substance abuse measures.

So the process for discussion and voting. So thank you for all of you who reached out to us in advance to let us know you would not be able to attend both meetings, or one or the other. So because of the participation numbers, we will not be voting today during the call, but we will be following up the call with a recording of both meetings, as well as a link so that you can do your vote via survey.

So the way the meeting will be conducted today is we'll do a brief introduction of each measures. The developers will do that right before we start our discussion. Our lead discussants will begin the committee discussion by criterion, and then the developers will be able to respond to each of those questions at the discretion of the committee. And then the full committee will discuss and you will be able to vote via SurveyMonkey following the meeting.

So our first major criterion of endorsement is importance to measuring reports. And so course that includes our evidence and performance gaps. There's a decreased emphasis on evidence, and we require the measure developers to attest the evidence doesn't change, evidence from the last evaluation and the standing committee to affirm if there is no change in evidence. For a performance gap, there's an increased evidence on performance gap, and data on the current performance gap in care and value - and variations.

Our second criterion is scientific acceptability, which includes reliability and validity. And so there has been no difference in the measure specifications.

We want them to be precise and with information needed to implement the measure. And so we require updated specifications.

And for the reliability and validity, including risk adjustment, there's a decreased emphasis if prior testing is adequate, additional testing is not required, unless there has been a change in the data source, or the level of analysis or setting, if the previous testing was limited to - or if the previous tested was limited to face validity only. So all measures must address the use of social risk factors and risk adjustment approach.

So our third and fourth criterion are feasibility and usability of use. So for feasibility, it's making sure the measure is feasible, including the e-measure feasibility assessment. And with usability in use, we will - we - you will be voting on them separately after the meeting via survey. But use is used to make sure that it's used in an accountability application and public reporting. And for usability, we're looking at the impact and unintended consequences.

So as I mentioned, we will not be voting today during the call because we do not have 66% of the committee. But when you do vote, the team will be compiling the voting results from the survey, making sure that we have quorum, and making sure that if there are any measures that are consensus not reached, we will be able to reach out to the committee, and we will have to revote on that during our post meeting call.

I will pause. I know that was a lot of information, even though it's familiar. But I'll just ask what the changes, so if anyone has any questions?

Jeff Susman: This is Jeff Susman. The only comment I'll make is that when we don't vote online, it means I have to go back and sort of remember what my thoughts and make sure that, you know, I'm accurately voting. So I don't know if there's

any opportune to be able to vote for those who are online attending currently, and then follow up with the voting for the people who couldn't make the conference.

Desmirra Quinnonez: Yes. Thank you.

Jeff Susman: It just makes it pretty clunky and we duplicate effort.

Desmirra Quinnonez: I do understand that, and I'm sorry about that, Jeff. We do - we would like to do that. However, what we'll be doing - we do recommend that you take notes during the call. We will be sending the recording, but we'll also - the transcript is available 24 hours following the call. So we will be following up with the transcript. So that might land another assistance in reminding you of what the conversation was during the meeting.

Jeff Susman: Yes. I mean, a transcript or listening to the recording, you know, frankly, I'm just not going to do it. I will make some notes to remember as I can. It's just - it's another step that I think adds to the burden for the committees, and I don't see any reason from a sort of accuracy perspective why we need to separate out those. But it is what it is and I'll make do, for sure.

Michael Abrams: Yes. This is Michael here. Thank you. So thanks for making the comment. I mean, we did think about that. The one thing we will try to do today if we can, given that the voting doesn't happen half - doesn't have to happen online, is to give you a bit of time back.

But for technical reasons, we thought it would made sense to just make a singular survey as opposed to doing a separate SurveyMonkey. That's why we chose to do that. So as a small remuneration back to you, we'll try to give you a bit back time - time back because of that.

Jeff Susman: Okay. That sounds fair. Thank you.

Michael Abrams: Yes. You bet.

Michael Trangle: This is Michael Trangle. I know that we're stuck with this for this time, but I'd like to sort of share that sentiment, knowing how much better my memory is, if it's more recent in my mind than trying to recall the conversations from a while ago.

Desmirra Quinnonez: We completely understand. Thank you for your comments. All right. With that, we will move forward. I will turn things over to Peter so we can begin the measure discussion of 0560.

Peter Briss: Thank you for that. So 0560. I have the primary discussion, discussion being David and Lisa, if either of you would like to start, that would be great.

Lisa Jensen: Hi. This is (unintelligible).

Michael Abrams: Just a moment. This is Michael. Sorry, Peter. Let me suggest that we have - see if Stewart is on the line and maybe they progress first, and then we go into having the discussions, okay?

Peter Briss: (Unintelligible).

Michael Abrams: Is somebody from the Joint Commission on then and who - could you identify yourself and give us a brief overview of your measure targeting, what you know about the comments thus far?

Elvira Ryan: Hello. Yes. This is Elvira Ryan from the Joint Commission, and I am the clinical lead representing these measures today. This is our HBIPS-5 measure, which is the appropriate justification for multiple antipsychotic medications. This measure was implemented by the Joint Commission in 2011. It has been NQF endorsed since 2009.

The patient population for this measure are patients discharged from inpatient psychiatric facility. The age range is over a year old and upward. This measure requires that patients who are discharged with two or more antipsychotic medications, have documentation of appropriate justification.

The appropriate justification is specified as a history of a minimum of three failed multiple trials of monotherapy, documentation that if the patient is discharged on more than one medication, there is a recommended plan to taper to monotherapy, or there's documentation of a plan for across taper being in progress.

The other documented justification is if the patient is being medicated for augmentation of clozapine, and those are the three that were the evidenced based justifications that will allow the case to pass the measure. There is a provision that if there's documentation of another justification, that can be recorded and abstracted. However, that will not pass the measure.

For the Joint Commission, this measure is required for accreditation by the freestanding psychiatric hospitals, and it is optional for psychiatric units within an acute care facility. This measure is also in the inpatient psychiatric facility quality reporting program for CMS. And CMS requires that all freestanding psych hospitals and psychiatric units paid under the inpatient psychiatric facility prospective payment system, report this measure. Thank you.

Desmirra Quinnonez: Thank you.

Michael Abrams: Thank you.

Peter Briss: Thanks very much. And now, I'll try to keep from getting out in front of myself again. David and Lisa are the primary discussants. And so we'll try to work sort of one question at a time. So first, any comments on evidence please.

Lisa Jensen: Hi. So this is Lisa. And in general, the evidence presented from the previous review was judged to be high in quantity, quality and consistency. And they did do an updated literature review, but there weren't really any new guidelines or anything that would warrant a change. And in general, felt that there were no new studies, that it was still relevant, the literature that was there.

And one sort of thing was that there didn't seem to be any studies looking at one of the commenters as a percentage of patients who were actually being discharged on multi antipsychotic in terms of how prevalent this was.

Peter Briss: Thanks, Lisa. David, anything to add.

David Einzig: Yes. So just - I'll definitely (unintelligible) on a couple of things. So this measures, talking about patients discharge on multi antipsychotics was not specific obtained diagnoses. This is based off of guidelines from 2004, specific for schizophrenia. Correct me if I'm wrong, but we have guidelines for force schizophrenia.

And so this measure is not specific for schizophrenia, but rather anybody

discharged on multiple antipsychotics. So that may include folks with bipolar disorder, autism or fetal alcohol or other neural developmental conditions. So that'd be my main chime in on that.

Peter Briss: Thanks. And with that, anybody - would anybody else like to make comments on the evidence criterion?

Harold Pincus: Hi. This is Harold, and I was one of the reviewers, and I have a lot of concern about the evidence criterion and the updating of it. I go along with what David said about something (unintelligible) other concerns as well. But if you look, most of the reports that we're reading and papers that are reviewed are over 10 years old.

And in fact, there's been some recent evidence, including a paper from Tiihonen and others in JAMA psychiatry this year that found that - that looked at the association of antipsychotic polypharmacy versus monotherapy, and found that with regard to psychiatric re-hospitalization and also schizophrenia, and the findings of the study suggest that certain types of polypharmacy may be associated with fewer re-hospitalizations and monotherapy.

So the evidence space has gotten a bit more crowded in recent times. And so I'd be concerned about the evidence based, you know, as is - measure being based upon evidence from such a long time ago.

Peter Briss: Thank you. Other comment from the committee?

Julie: This is Julie. I have to agree that if we follow the evidence and we have new evidence, given the specification of the measure, three failed monotherapy attempts, et cetera, that inadvertently against the current evidence, or at least without sufficient evidence that it could no longer be valid, right? People

would be dinged, if you will, for something that was more appropriate than three failed attempts at a single antipsychotic. So I have to agree with what was said.

Constance Horgan: This is Connie. I just wanted to point out that this is just - it's unfortunate that the APA panel on guidelines has not yet met, but that is planned for the fall, if I'm not mistaken. Perhaps someone has more information on that. Are new guidelines forthcoming?

Peter Briss: Does anybody have a specific response to that question from the American Psychiatric Association?

Constance Horgan: It was presented in the evidence. It was presented by the developers that a new panel coming is - will be meeting starting, I'm not sure, I think in the fall, on guidelines on this. Was that not presented in the evidence in the readings by the Joint Commission?

Peter Briss: So could the developer answer that question please?

Elvira Ryan: Hello. This is Elvira Ryan. At the time we submitted the submission, there was nothing on the Web site with a timeframe. However, I looked at the Web site again this morning for the APA, and it indicates that they expect something to be happening in September of this year.

Peter Briss: Thank you. Are there …

Jeff Susman: I think this is - this is Jeff Susman. I think this is one where, because the growing controversy in this arena, that this measure unfortunately no longer meets our criteria for good evidence and to hold it up as a quality metric may be doing patients a disservice. So I would vote against that on evidence.

Lisa Jensen: This is - I'm sorry, this is Lisa. And I wasn't aware of the article, but I just pulled it up and I mean, I guess it bears looking into more. But basically the conclusion was that clozapine with aripiprazole was the best combination, and that aside from injectables or clozapine, that when you compare other two antipsychotic combinations, there didn't seem to be as strict superiority demonstrated from the combination versus monotherapy in the individuals with schizophrenia that were looked at. So I just wanted to put that out there.

Tami Park: Yes. This is Tami. I mean, I've been doing a quick search too, and there's also a 2014 schizophrenia research article that finds that in general monotherapy versus polytherapy, polytherapy is - leads to more mortality and psychiatric hospitalization. So sounds like we're saying that a more rigorous mental scan needs to be conducted.

Peter Briss: So it sounds to me like there's a significant concern about the level of current evidence to support this measure and at least more work needs to be done. The process that we're trying to do today makes the conversation a little bit harder than it would normally be, because evidence is a must pass criteria, right? And so we would - criterion, so we would not - were we voting now, we might vote down the measure on evidence and then we would stop and move on.

So I've heard a number of people expressing concerns so far. I'm a little tempted to stop discussing this one at this point and just see how the voting plays out on the evidence and discussing this one further in a follow up meeting if it's necessary. Does anybody feel like - does anybody feel strongly that they want to continue the discussion on the other criteria at this moment.

Tami Park: No.

Lisa Jensen: I'll go along with the committee just to get more information on the evidence, but honestly I am not convinced that there's great evidence to show that polypharmacy is better than monotherapy, other than the conditions that are already part of the measure. But I'll go along with what the group says.

Dolores Kelleher: Great, and I think that's not so much the point for me. This is Dodi. As much it is that there hasn't been an updated scan of all the available evidence so that it supports or doesn't support the current measure.

David Einzig: And this is David Einzig. From a clinical perspective, again just to emphasize, my concerns has to do with guidelines being specific for schizophrenia, when this measure has to do with polypharmacy, not specific to condition. And second, current evidence regarding the current appropriate justifications of three failed trials of monotherapy or cross tapering or augmentation with clozapine, I believe there's some evidence out there for low dose augmentation with the Abilify for elevated prolactin associated with (Respinal).

And then in other populations, again clinically speaking, not for schizophrenia, but it does make clinical sense in real situations that don't go along with these specific outlined appropriate justifications where it does make clinical sense to do low doses of few antipsychotics, as opposed to high doses of one antipsychotic, and clinically it does the patient good and there's no harm.

Peter Briss: Okay. So at this point, I'd like us to - unless anybody - I've heard many people agree with stopping the discussion now and voting on the evidence. And we'll talk more about this measure at the follow up meeting if we need

to. I heard a number of mostly yeses to that. Does anybody else want to object to that process?

Okay. Hearing none, staff, can you make sure that we're voting just on evidence when you send out the voting link and be clear about what the question is at this point please.

Michael Abrams: Yes. That is - this is Michael here at NQF. That's perfect, Peter. You read our minds literally over the webinar here. So we will step that up and let's proceed to the next measure.

Peter Briss: Thank you.

Michael Trangle: I have a question. This is Michael Trangle. Is there going to be a summary of our discussion or not really?

Peter Briss: Michael, (unintelligible). There'll be a transcript tomorrow, right?

Michael Abrams: Yes.

Dolores Kelleher: Yes. We're going to ..

Michael Trangle: That'll be left available for us and the people that didn't participate to peruse?

Dolores Kelleher: Yes.

Michael Abrams: Exactly.

Dolores Kelleher: So following the meeting, we will send the recording of the meeting, as well as the survey link and then we will send the transcript as well. It's available

24 hours after the call. So we'll make sure that we send that right away as soon as we receive it.

Michael Trangle: Okay. Thank you.

Dolores Kelleher: You're welcome.

Peter Briss: All right. So any other comments anybody wants to make before we move to the next measure? Hearing none, the next one is 0640, hours of physical restraint use. And this time I won't go out of order. If the Joint Commission would like to tee this up for us, we'd appreciate it. Thank you.

Elvira Ryan: Yes. Hi. This is Elvira Ryan again. This measure looks at the hours of physical restraint use, as it's termed an event measure. This measure was implemented by the Joint Commission in 2011. It was endorsed by NQF in 2010. The patient population for this measure again are all patients that were inpatients in a psychiatric unit or a psychiatric facility.

This measure applies to any patients above the age of one. And this measure looks at any patient who was restrained during their hospitalization, and it looks to determine the hours of restraint used per 1,000 patient days. It is used by the Joint Commission for accreditation of freestanding psychiatric hospitals. They're optional for psychiatric units within an acute care facility.

This measure is also in the CMS inpatient psychiatric facility quality reporting program. It applies to their freestanding psych hospitals and units paid under their prospective payment systems. Thank you.

Peter Briss: Thank you. And so I have the primary discussions on this one as Brooke and Bernadette. And I believe Bernadette is not here today. So with that, Brooke, can you tee this up for us?

Brooke Parish: Good morning. Yes. Again, this is a measure that is sort of an older measure coming back, and we're taking a peek at it again. In terms of evidence, the concern that it is evidence - certainly has shown that the concern of is for minorities and elderly being overly strained is not necessarily the case. In fact, if you are a young white male, that is most likely.

Certainly, in terms of new evidence, there really were no new studies. And in terms of performance gap, again the concern that the older population is being restrained or minorities outside of Caucasians again is also not necessarily the case. The good news is that you can see that there has been a decrease every year in pretty much the use of restraints.

Peter Briss: Thank you. Any other comments from the committee on the evidence criteria?

Michael Abrams: So this is …

Tami Park: Where is the data showing the decrease?

Brooke Parish: It's coming from the hospitals.

Tami Park: I mean, is that …

Michael Abrams: So Tami, so Michael here at NQF. So my reading and maybe the - Elvira can comment about this. You did - you presented a person random effect model that actually showed an increase in rates over time, which was different than

the citation that you gave on Rashinki - the Rashinki study which showed decreases. Maybe you want to comment? Maybe it would be useful for you, Elvira, to comment about that sort of ironic results presentation or findings presentation with time that you presented.

Peter Briss: So that's a question for the developer. Can you comment please?

Dave Morgan: Yes. Hi. This is Dave Morgan. I'm a statistician with the Joint Commission. And the paper we cited used a baseline period as a comparator for - it was a one year see …

Man 1: Five years.

Dave Morgan: Five years. But for the random effects model, I just used a continuous variable and looked at the slope of that. So it was a different methodology.

Jeff Susman: Is there any reason - this is Jeff Susman. Is there any reason why using a different methodology would be expected to give such a different result?

Steve: Yes. This is Steve (unintelligible), statistician. The population over time changed, and that could have an effect on looking at the longitudinal trend. Whereas looking at cohorts, you're looking at a specific group of people over time, but you're basically controlling for that group of people. Whereas if you're looking just over time without considering cohorts, then the change in the population over time could have an effect on that overall trend.

Jeff Susman: Thank you.

Leslie Zun: So this is Leslie Zun. I have a couple of questions related to this measure. So one, the way I understand this, it’s just to measure the time that inpatient

psychiatric patients, not all patients that get restrained in a hospital environment. So that's one. And then two, I wanted to take this a step back because I'm trying to figure out the evidence that correlates the use of restraint and or seclusion to - for outcomes.

Tami Park: I also had that question.

Leslie Zun: Because I've actually done research in this area, and in my own study in this area - and I'd like to see the evidence that shows that correlation, because if we're going to put hospitals through this, we've got to have a purpose.

Peter Briss: So would the developer like to comment on that issue please?

Elvira Ryan: As noted earlier, this measure has been implemented for quite some time. And at the time of implementation, the focus was looking on the quality of patient care and patient safety, and the hope of decreasing the risk of patient harm because of the use of restraints and seclusion. So that has been the primary focus. The measure just looks at the time that the patients are in restraint. The measure does not look at patients' outcome.

Leslie Zun: This is Les again. Yes, I understand that a number of changes occurred with the new policies and procedures put forth by CMS and The Joint Commission. So the question is still, is it a useful measurement at this point in time? And I don't - it doesn't sound like we have an answer to that.

Michael Abrams: Any other comments around the table please?

Lisa Jensen: I did notice that the …

Michael Trangle: This is …

Lisa Jensen: Oh, pardon.

Michael Trangle: (Unintelligible). I'll go after.

Lisa Jensen: I was going to just say like in the beginning or intro of the measure, it does state in terms of decreased costs of care and other such comments that I did not find necessarily supported in the literature or in the evidence presented.

Michael Trangle: This is Michael Trangle. I think clinically there is some correlation, but what's causing what effect is not clear to me, you know. You could say people are sort of getting out of control and they end up in restraints, you know, with whatever degree of injury for sure.

You could also say the fact of trying to put somebody in restraints causes a reaction and you have more injuries and stuff. So there is some risk factor that's real there, but what's causing what effect is not clear to me.

Tami Park: I mean - this is Tami. Yes, on that question, I've heard anecdotally from clinicians working in psychiatric units that this was an important measure, that it led them to implement training around de-escalation and stop them from - stop the staff using restraint as frequently as they had been. That's just anecdotal. I mean, if the developer had similar evidence, maybe that would be helpful in trying to think about this.

Jeff Susman: Yes. I mean, certainly - this is Jeff Susman. The practice of restraints for elders and the geriatrician have dramatically changed. And I think there's reasonable evidence, although not necessarily applicable to this population, that this is a positive. The question for me is probably one of usability at this

juncture and what have we gotten the most bang for our buck and, you know, in the sake of parsimony in measurement, maybe this is one that at this time.

Michael Abrams: Yes. So let's keep it - let's hold on - Michael here at NQF. Hold on new future criteria. Stay focused on evidence, whether or not there's more discussion or questions that are related to whether or not this measure is connected to outcomes that are desirable. The question about evidence that we're focused on here. So I just want to keep us focused on that, supposed to jumping to usability or anything else. Yes.

Peter Briss: So this is Peter again. Does anybody else have comments that haven't been made yet sort of the evidence criterion? All right. I think I hear sort of more diversity of opinion on this one than on the last one. So I think I'd like to go on to the other criteria and think about performance gap please. And Brooke, can you walk us through that please?

Brooke Parish: Certainly. As you can see, there's a number of hospitals that are reporting this evidence or reporting data. Again overall, you will see a running decrease in terms of the amount. Interestingly enough, there are several studies that they cited in terms that if you were African-American male or a minority male that you would be at greater risk of restraints or seclusion.

However, the data from the hospitals appear that that is absolutely not the case, and in fact it is the Caucasian young male. And again, as I mentioned before, that it was not - like from the data from the hospital, it does not appear to be the elderly again.

Peter Briss: Comments from the - yes, I’m sorry. Comments from the committee?

Leslie Zun: So this is Les Zun again. So if you - excuse me. If you look at the requirement which says quality improvement, is this - is the - the issue of quality improvement, is that based on, for this part of the measure, the measurement, does it have to have a quality improvement that is proven in order to apply this measure - this data?

So if you keep going up and up, right there, 1B. Performance gap requirement includes demonstrating quality problems and an opportunity for improvement. Do we have to prove that? Because what this tells me is, it gives me the data, okay, but it doesn't tell me - it's not demonstrating quality or opportunity for improvement. That's where I'm confused.

((Crosstalk))

Michael Abrams: Yes. Oh, go ahead, Peter.

Peter Briss: Yes. This is Peter. So the - it's probably the - whether the measure has been - has done what we hope that measure do and generate real improvement in the real world is generally something that we consider in the usability and new criterion later. So we might want to table that discussion.

And then just parenthetically, I'll say that we and other committees have approved lots of measures where in spite of the fact that we haven't been able to show as much improvement as one might wish for, I do think - I wanted to ask folks whether - I don't have a good sense, but an hour of restraints used per whatever the - per 1,000 inpatient hours, sounds to me like it - that might be arguably topped out.

And if somebody who's closer to this problem than I am wants to comment on

whether this might be or might not be considered a topped out measure, I'd love to hear perspectives.

Tami Park: I think the hour - this is Tami. I think the hour issue is - the way it's presented as hours makes it confusing. I was confused because it's such a big denominator. I know why they do that, because they want control for the time that the person is in the hospital. So what would be more helpful or would be helpful as well is to see how many patients this refers to. And I don't know if that's what number of cases on the bottom of that table means.

Peter Briss: That might be a question for the developers. Would a developer like to comment on that issue?

Tami Park: Somebody?

Steve: (Unintelligible) denominated.

Dave Morgan: So the number of cases is numerator.

Michael Abrams: Oh, yes. Number of patients with restraint.

Tami Park: So for example, in 2018 …

((Crosstalk))

Tami Park: … 36,000 patients, 36,192 patients were restrained?

Dave Morgan: Yes.

((Crosstalk))

Dave Morgan: So I think it's just the number of occurrences. It's the number of cases. A patient could be restrained multiple times.

Tami Park: Right. Okay. But it's a lot of incidences.

Dave Morgan: Correct. Yes.

Peter Briss: Okay. Thank you for that help. And anybody else have comments that haven't been made on performance yet? Hearing none, Brooke, back to you on reliability please.

Brooke Parish: I mean, the main issue here with reliability is that we are looking at freestanding psych hospitals, most of the freestanding psych hospitals do not use a EMR EHR. This is a human recorded measure. It is requiring somebody to know exactly when somebody went into restraints and when somebody went out of restraints and keep a very accurate record of this.

One of the concerns I did have is probably the hospitals that do restraints the least, are probably the ones that are - do not have a time keeper of exactly going okay, they went in at 6:07. And the hindsight of while they're doing their documentation several minutes later would be well, it may have been around 6 o'clock versus the hospitals that have - actually have a scheduled time keeper. So again, it is a human required or acquired measure.

Peter Briss: Anybody else have comments?

Jeff Susman: On inpatient units, it is human required, even with an EMR. Someone has to pay attention. Someone has to start the clock. Someone has to look again. And you can see when it's written, but it's always sort of like a human thing.

And operationally, whether you have an EMR or not, it's a tough thing to work on. And if it's important, it's worth it. If it's really not so meaningful, it's a big deal, you know.

Peter Briss: And just in terms of reliability, I think that they - that the Joint Commission at least did re-abstractions by Joint Commission staff for we record some number of people, like 190 and shows apparently perfect agreement. So there's at least that reliability piece of information.

Michael Abrams: Right. So Peter, let me piggyback on that comment and for the previous comment. This is Michael at NQF. So the reliability that was done, and perhaps it will be useful for the developer to comment about this, was just as you described.

It was just abstracting and re-abstracting the record. It wasn't like there were two people standing by and watching the incident with two different stopwatches and recording the exact time, which would arguably be a more precise type of reliability.

That isn't possible with the data that they have from their EHR. But if the developer could just comment about what their expectations were about abstracting and re-abstracting. You've got perfect agreement. Do they really - how strongly do they feel that that reflects - and they expected to see much variability from that at the time?

Elvira Ryan: At the time that the measure was tested, it is correct that the Joint Commission traveled to the facilities and looked at all of the records that the hospital abstracted, and we were looking to see that the abstraction was a match and that it was accurate. We did not go to the site and observe patients being

placed in restraints and then look at the documentation that occurred from that perspective.

At the time we do the testing, we would always like more cases than we can get. However, our testing is time constricted and this is the data that we had available from the original testing.

Peter Briss: And clearly nobody is ever going to do a reliability test with sort of an outside observer doing a time motion study with a stopwatch. So there's that. Anybody else have comments that haven't been made that they'd like to get into the record on the reliability criterion please?

Jeff Susman: This is Jeff Susman. When you said reliability, were you looking at the reliability of the number of hours abstracted, or just whether someone was restrained or not?

Peter Briss: So could the developer comment on that please?

Elvira Ryan: It was both.

Jeff Susman: And they were both 100% reliable?

Elvira Ryan: Go through the testing.

Jeff Susman: (Unintelligible) to be honest, knowing the settings, but okay, if that's what it was.

Peter Briss: And anybody else have comments or questions that haven't been put on the table yet? Hearing none, let's go back to Brooke for teeing up validity please.

Brooke Parish: Certainly. In terms of validity, it does seem to have a fairly good face validity, an empirical validity. There is a HBIPS-3 measure that does show a correlation to, and that is three hours in restraints. Overall, it is considered to be moderate rating.

Peter Briss: And isn't HBIPS-3 seclusion?

Brooke Parish: I believe so, yes. Seclusion.

Peter Briss: Yes. Okay.

Michael Abrams: It is. Yes.

Peter Briss: There's some modest or moderate correlation with another measure that they thought might be related.

Michael Abrams: Point - Michael at NQF. Point 26 was the correlation with seclusion use. So monitor is a good descriptor.

Peter Briss: Anybody else have comments or questions about this one?

Leslie Zun: So this is Les. So as far as the measurement is concerned, were institutions allowed to use whatever measurement tool they wanted? Because I'm not sure that there had to be a timekeeper there. For instance, in one of the institutions I worked, it was - the time of restraint was the time the provider put the order in, versus someone looking or a nurse putting in or estimating the time. I think there's some variability in how institutions were obtaining this information.

Peter Briss: Would a developer like to comment on that issue?

Elvira Ryan: The specifications are written such that the time that is to be abstracted is the actual time that the patient is placed in restraints, and the time that the patient is taken out of restraints. And then they are to compute minutes and convert those to hours. So then that is calculated in terms of the number of hours per patient days. So the specifications require that it is the actual time for the patient being in the restraint.

Peter Briss: And then - and Les, at least for the number of people dual abstractions they did, again they got perfect agreements.

Leslie Zun: Okay. Thank you.

Peter Briss: Anybody else have new comments or questions on the validity criterion? Hearing none, let's move to feasibility please.

Brooke Parish: Certainly. Feasibility, again this is a maintenance measure, so it has been done again. And I think most of the concerns within that feasibility would go back up more towards reliability that have already been discussed. The preliminary rating on feasibility has been moderate.

Peter Briss: Additional comments or questions?

Lisa Shea: This is Lisa Shea. I just did want to point out in terms of keeping time, there's a lot of other regulations that do require time be kept for when an order is gotten, when it’s going to be extended, et cetera, et cetera. So at any rate, institutions have to keep time because of all the other rules that are - come with restraint seclusion.

Peter Briss: Thank you. And anybody else, comments on feasibility? Hearing none, back to usability and use please.

Brooke Parish: Again for the usability, this is a maintenance measure. It has been done. There is typically accountability. This is not something - seclusion and restraint is not done in sort of a vacuum, but more as a team effort. And the preliminary rating for use have been a pass.

Peter Briss: And any additional comment or concerns?

Brooke Parish: Also under usability, there was one concern of looking - if you are working at trying to get seclusion and restraint, and this is the restraint measure down, whether you are increasing the use of PRNs and putting people more at risk for side effect from the PRN medication.

Peter Briss: Thank you. Any other comments are concerns on usability and use? Hearing no …

Leslie Zun: This is - sorry. This is Les.

Peter Briss: I'm sorry. Go ahead.

Leslie Zun: So I question the cause versus the effect on this. So is the restraint causing more - I mean, is the approaching the patient causing more problems or less problems? I mean, just I'm not sure that restraint tells us that there's inadequate or inappropriate care.

Peter Briss: So that's probably - and this - I think we teed up some of that in the evidence discussion. And so for people that - the way which is where I would probably put it, if you feel that the - so if any of the committee members felt that they

were concerned about the likely tradeoff between benefits and risks of applying this measure, you might vote it down on the evidence criterion if you're convinced that the data presented suggests a net - scientifically documented net benefit, ten you would probably vote up on the evidence for that measure.

Brooke Parish: Lisa, do you have any thoughts on those evidence on this? Interested in hearing your thoughts.

Lisa Shea: Which thoughts are you looking for?

Brooke Parish: Lisa Shea on the use of restraint. Like is there evidence - some evidence that it's harming the patient?

Lisa Shea: Sure. I have anecdotal - I mean, just from my own experience working on an inpatients gastric unit is that when this measure came out, I mean I think at least at the hospital that I previously worked at, it made a complete culture change and really impacted the reduction of restraint seclusion in such a positive way, and, you know, staff satisfaction improved, patient experience improved, injuries went down.

I think that institutions really take it seriously, and it's I think overall made a very positive impact. That's not to say that people are restrained or secluded, that the institution is doing anything wrong, but at least now I think that there's a lot of things put in place to try to help patients so that this doesn't happen to them unless it has to, and it's a lot less traumatizing for everyone.

Lisa Jensen: This is Lisa Jensen, and I would just like to comment also as a nursing representatives here and chime in with what Lisa Shea just said. I think the

reduction - the goal of reducing the use of both seclusion and restraint is always forefront for nursing on the inpatient units.

I know here in the VA, we really strive to reduce the amount of time that we have to use either of those interventions to deal with patients' behavioral issues. So I think it's an important thing to look at as well. Thanks.

Peter Briss: Thank you. So with - I think we heard both - likely both poles of that discussion and the evidence discussion and we've now heard them again here. So any more comments on - first, if anybody feels like they want to comment again on the evidence point that - on a point that hasn't already been made?

Morgan Shields: Hello. Can you hear me?

Peter Briss: Yes.

Morgan Shields: Hi. My name is Morgan Shields. I'm a PhD student and I'm (unintelligible) just from like public group. And this has not been published, but I just feel compelled to share that I've done some preliminary analysis looking at the effect of the IPFQR and intervention on use of restraints (unintelligible), and there is evidence that it reduce duration of both restraint and seclusion.

Peter Briss: Okay. Is that a public comment? Are you on the phone as a member of the public?

Morgan Shields: Yes.

Peter Briss: Okay. Thank you. Anybody else have comments on usability and use then? Hearing none, Brooke, anything you want to say about relating to the competing measure?

Brooke Parish: There are a couple of related. One is restraint prevalence and it only is 18 and up, and only does vest and limb. And then there's for long term nursing or long term care, percentage of residents who are physically restrained. Also mentioned was the other HBIPS that does seclusion, which could be considered also related.

Peter Briss: Okay. Anybody else, comments or questions about related and competing measures? Hearing none, anybody else have a closing comment that hasn't already been made? All right. So I hope everybody took good notes and we'll see how the post meeting survey goes. And if nobody else has further comments, we'll move to 641, hours of seclusion. And I have Bonnie Zima and Andrew Sperling as the primary discussants.

Bonnie Zima: Andrew, do you want to take the lead?

Andrew Sperling: So I'll be brief (unintelligible). So this is - the measures around use of - just like with restraints, measures of seclusion have been around for quite some time. As it was mentioned earlier, this use and measuring this, you know, began - efforts to deal with this began as far back as 20 years ago, when there were a number of exposés in various newspapers about bad outcomes associated with use of restraint seclusion.

And the measures related to this began as conditions of participation imposed by then HCFA, now CMS, as far back as the early 2000s. and it really has brought about a change in culture in inpatient psychiatric settings of how this is used, both seclusion and restraint, and how it can be reduced and how we actually produce better outcomes when seclusion restraining orders (unintelligible) earlier, it's brought about a change in culture of the staff at work and inpatient psychiatric facilities. So this is not from nominee

perspective. This is not anything new that we need to measure this and we need to do - take steps to limit its use.

I would comment that we still struggle with a number of things around the cause and effect of this. It is the case that individuals who have to go through typically (unintelligible), but also seclusion as well, sometimes patients actually volunteer for seclusion because they feel threatened by another patient in the inpatient unit.

Sometimes it's actually done in a punitive way and we all know obviously we want to limit that so that when it's imposed on a patient by the staff in a punitive fashion. The measure doesn't really pick that up, but we certainly want to limit the frequency of it, and we've had a number of measures out there from various accrediting agency for quite some time. So I view this as somewhat mature in the field in terms of how long we've been measuring imposition of seclusion.

Peter Briss: Thank you. Oh, go ahead.

Andrew Sperling: Thank you. That's it.

Peter Briss: Thank you. And I keep trying to go out of order today. I apologize for that. Would the Joint Commission like to add any additional teeing up for this measure please?

Elvira Ryan: Yes. This measure is very similar to the restraint measure. It was implemented by the Joint Commission in 2011, endorsed since 2010. It is the same population, patients discharged from psychiatric unit or a facility. It is the same age group, calculated the same way, looking at hours of seclusion per 1,000 patient days.

It's used in the accreditations for the freestanding psych hospitals. And it is also used and required for the CMS IPFQR program. Thank you.

Peter Briss: Thank you. And Bonnie, would you like to add anything on the evidence criterion for this one?

Bonnie Zima: Yes. I think that actually Andrew's perspective actually brought in some important new information, and we need to be mindful that this is - has changed the culture on inpatient units. And also to echo my colleague in nursing that this remains a priority.

However, on the evidence, again the same arguments against evidence continue, just like the measure we just talked about. There's no updates provided. Evidence is not graded. You know, there's no evidence documenting that adherences to this measure relates to meaningful outcomes.

Echoing Lisa's point earlier about cause and effect and the problems of the direction of the relationship, whether you adhere or not is not really adequately teased out in the literature.

Peter Briss: Thank you. And yes, much of the pulse of that discussion sounds just like the last measure. And anybody else like to comment on the evidence criterion?

Michael Trangle: This is Michael Trangle. Yes. I think that the evidence is actually less present for seclusion and for restraint. I think the general agreement in the field that restraints are a risky business, and you really want to deescalate verbally and sometimes with medication to get people calm down in their own resources, as well as some ancillary resources.

Seclusion sometimes is thought of going seclusion, we can prevent a restraint, and it's not necessarily, at least clinically, always viewed as a bad thing, especially if the person is saying, I need a time out. You know, if it's locked, if a seclusion is unlocked, it's not technically.

But I guess the point I'm making is it's - the evidence is less and the clinical impact is also less. It's almost something along the continuum, than something we know is bad clinically by itself.

Peter Briss: Thank you. That's an important perspective. Anybody else like to comment on the evidence on this one? All right. Hearing none, I'd like to go ahead and move us to performance gap please. Andrew, do you want to tee us up?

Andrew Sperling: For performance task?

Peter Briss: Yes, please.

Andrew Sperling: So I think what was most startling to me in the submission was the racial disparities.

Bonnie Zima: And this is Bonnie.

Andrew Sperling: It seemed to be very different than the restraint, which was - I'm curious as to why that's happening. But the way - the disparity along racial lines was somewhat disturbing.

Bonnie Zima: And this is Bonnie. Just as far as the performance gap on this, the developer reports, as for performance data over a 10 year period within a mean that's steady between point 26 and point 275. But there's also some outliers in 2015

at point 968, 2016 point 576, and outlier year 2017 at point 358. But the question there is, are these meaningful differences? I honestly couldn't tell.

As far as offering any evidence of improvement over time, the developer cites a study supporting the effectiveness of a restraint and seclusion reduction program in one hospital. And as Andrew suggested, yes, there was variation by associate demographics. But it does not address for clinical severity, and I think this is probably because of the limitations of their data source.

Peter Briss: Thank you. Anybody else have comments on performance gaps? Hearing none, then I move us to reliability.

Bonnie Zima: Andrew?

Andrew Sperling: Well, I would just note that - I mean, there are a number of other accrediting agencies out there that have been assessing this for more than a few years, actually going back more than a decade in many cases. So I view this as very reliable.

I didn't actually think about it until the discussion earlier on the restraint side, the timing issue and the - if some facilities actually have staff with stopwatches in the unit and others don't, that could raise some concerns about reliability. But I view this as again somewhat of a mature measure that other accrediting agencies have been assessing facilities on for a number of years.

Peter Briss: Bonnie?

Bonnie Zima: Yes. And from my perspective, I thought the specifications were clearly specified. But again, back to the argument on the prior measure of two hours of physical restraint use, it looks to me, and I think what I heard from the

developer today, was that the findings on the 100% agreement across two chart abstractors was done during the development or early phase of this measure.

There was 190 cases reviews that hospitals had abstracted. And the way I read it was that of the 190, only nine patients had one seclusion, and that the three (depend) variables that they looked at for agreement were simply the date, the event date, the event type and the number of minutes. And so in those nine cases, there was 100% agreement across both chart abstractors. So the question is, is this is sufficient to continue to rate as moderate?

Peter Briss: Okay. Anybody else? Hearing none, I will move to validity please.

Bonnie Zima: Andrew?

Andrew Sperling: Strong validity on this in my view. But again, this is - sorry to repeat myself here, but I've been doing this for over a decade in many cases and view this as sort of standard procedure in terms of how …

Peter Briss: Andrew, you're sort of - we partly lost you there in the middle of that sentence.

Andrew Sperling: I got cut off there a second. I think the facilities have been doing this and recording this for a long time, and I view this as something that they give you a routine business that they have to do every day. I don't see concerns around validity.

Peter Briss: Bonnie.

Bonnie Zima: Yes.

Michael Abrams: So Peter, Michael here in NQF. I'm making observations here. Andrew, I know you're a patient advocate at NAMI and everything like that, but I just want to clarify your language. When you say, it's important for you all as you're thinking about these measures and voting on these measures, when you say something is valid and you think it's valid based on your own experience, that's one thing and that's an important thing for you to consider. But is the data that's presented in the application showing you that validity? That's a separate question.

And, you know, they did one correlation here that was point 26. It's hard to call that a strong effect I think. It would be hard to find a statistician who would say that was a strong effect. So I just want to make sure that, you know, when we're discussing these things, that you're clear about, you know, your own experiences and why you think something is valid versus the validity presentation actually in the application, okay? Just a comment.

Andrew Sperling: I totally get your point and you're exactly right. I'm not a statistician. In fact, I remember years ago taking probability and statistics in college and I think I got a C minus. So you should not rely on my strength as a statistician.

Michael Abrams: All right. So Bonnie, do you want to make additional comments?

Bonnie Zima: Yes. So I think again similar to the other one, the developers' rationale for validity is a relatively weak but positive correlation between the two hour restraint use measure and this one at 0.216 value. So that's relatively weak, but you would expect there would be a little bit of a correlation.

The other thing too is again echoing the developer statistician reminding us that they had run a person regression modeling the seclusion of - with the total

patient hours using hospitals random effect, which makes sense. They said that there was significant improvement over time when they did this modeling using cohorts over time. But other than that, that's really it. There's evidence that adherences related to improved outcomes that are meaningful. Was I boring?

Peter Briss: No, perfect.

Michael Abrams: You're on point.

Tami Park: Yes. no, but I would just point out that all of the measures we look at have this kind of very disappointing limited validity tests where they take one measure that's been NQF endorsed and they correlate it against another measure. And I don't - I think they do that because they have to do something and it's expedient.

But I thought the discussion that we had before, I mean validity is whether the thing is basically measuring quality. Is it a valid indicator of quality? So I think the discussion we had before is really more to the point of validity than this statistical tech, which I agree is pretty light.

Bonnie Zima: And I agree with you, Tami. I think what happens in these situations, whether it's this or like an NCQA measure is that the developer is basically thumping her head against the limitations of their existing data sources.

Michael Abrams: And this is Michael at NQF. The other point that didn't come up, so I'll just remind everybody, there is a face validity presentation in this is well that especially at the standing committee, you are behavioral health experts, that is something for you to keep in mind, and that was favorable.

I'll just remind you what the stats were on that. Across their committee of 18 members, 13 said it had good face validity, five average and nobody stated that it was poor or very poor. So that adds to the validity story as well.

Peter Briss: All right. So anybody else have new comments that haven't been made on validity?

Bonnie Zima: Yes. This is Bonnie again. I think that looking at the original data on face validity, it was a total of 36 hospitals, and they were asked to rate their overall understanding of this measure. And 18 hospital reps had very good and 13 said good. Five was the average and zero was poor, very poor was zero. So they were rating the overall understanding of the measure and the clarity of the measure.

Michael Abrams: Yes. Thanks for correcting my numbers there. 18 were very good, 13 good, five average and no poor ratings.

Bonnie Zima: Those are 36 hospital reps.

Michael Abrams: Yes.

Peter Briss: Anybody else, comments that haven't been made?

Harold Pincus: Just in regard to the face validity - this is Harold. So the client and the face validity, Michael, maybe you could just say, isn't there an hour, a view that case validity is in - no longer meets the criteria?

Michael Abrams: Yes. So thanks for bringing that up, Harold. So one of the things this is relevant to is actually methods panel review. So as I said, you know, at the

beginning of this meeting, that quantitative review - and that committee doesn't have behavioral health experts for example on it.

So face validity there is especially something that is of concern as a singular sort of standard to see whether or not things look true or not. But at the standing committee level, we allow you to utilize your expertise more. Having said that, you know, the idea of a really good gold standard to be used in correlation and showing a strong correlation between a particular measure and that standard in a quantitative empirical way, is absolutely stronger than somebody saying it looks good and smells good, you know, as you're suggesting.

So I think to come up with a high rating, you would really want to have a good gold standard. But a moderate rating, a reasonable correlation with some other measures and then a face validity with a technical panel, still is something we see quite typically in NQF and that still often leads to a suggestion that endorsement is appropriate. Does that help you, Harold?

Harold Pincus: Yes. I just thought that there was some official decision that it will no longer be a consideration.

Peter Briss: I think, Harold, what you're thinking about - this is Peter. I think what you're thinking about it is, I think that there's a new guidance on when you know - when you're in the maintenance phase and you're not initially doing a measure of - they're now asking for some empirical validity testing when things come back in.

Harold Pincus: Right yes.

Peter Briss: It can't just be face validity. And so in this context, we're looking at a measure that's been around. It does have both face validity testing. And as is required, it has something else. We've - for both this measure and the last measure, there was some information on face validity and some information on, I would say mild or moderate empirical correlations with another measure that's supposed to be related. And so does anybody have additional comments on validity that haven't already been made?

Tami Park: This is Tami. I have a question. Where is the literature, the scientific literature comment? Because there's a fair amount of literature on these topics?

Peter Briss: So don't we have - don't we do most of that in evidence, right? a lot of what we're talking about is does the - what you do in the evidence criterion in kind of a simple minded way is, if you do with a thing that's being measured here, what's the evidence that people or patients would be better off, right?

And we've heard on each of the last two measures that - I hear some differences of opinion from the committee on the extent to which there's agreement about that. But that's sort of what one does in the evidence criterion, and that's where a lot of the existing literature comes in.

Tami Park: Right. Yes. For this - in this case, validity really is the measure as specified?

Peter Briss: Right. It's - and reliability and validity in this context are - reliability is, can you replicate the measurement and validity, it doesn't measure what it's supposed to be measured?

Michael Abrams: Yes. What it says it's measuring. Yes. Peter, your back and forth is right on point.

Peter Briss: We have done this a few times, Michael.

Michael Abrams: Good. Sure, indeed.

Peter Briss: Anybody else have new comments on validity? Let's …

Tami Park: And we have - I mean, the ultimate test would be if the - if you want to do like a randomized study, it's the hospitals that are implementing the measure, have better patient outcomes, including better patients and staff experience and the hospitals that are implementing the quality measure, which his beyond the capability of most measures at offer.

So we're stuck with these poor correlation. And I think we have to make a leap from the evidence from other studies to this measure, because we've never really going to have, or rarely have the evidence that we want specific to whether implementation of this quality measure leads to better patient outcomes.

Peter Briss: Yes.

Bonnie Zima: This is Bonnie. I think this discussion, just to flesh everything out, we haven't talked about the target analysis methodology that the Joint Commission has provided. It's a little different than what other developers have. I'm wondering if the statisticians want to come in on that quickly.

Elvira Ryan: Could you please repeat the question?

Bonnie Zima: The - so one thing that's a little unique in the validity argument is that the Joint Commission uses a target analysis methodology. So they're creating

their own kind of national benchmark, and then looking at the percentiles against that benchmark. Is that - can someone comment on that with what exactly they did?

Dave Morgan: Yes. This was kind of a guide for hospitals on our performance measure report. So the hospitals can see how they rate against how other hospitals reporting to the Joint Commission are doing. So it kind of gives them a range for which most of the data that we see are occurring, and then if they are beyond that range or not.

Elvira Ryan: Okay.

Peter Briss: Thank you. So the - I feel like we're circling a little on this topic. So does anybody else have anything else that hasn't already been said on validity? Could we please move to feasibility please?

Bonnie Zima: Andrew?

Andrew Sperling: I was muted. Okay. So in terms of feasibility, I feel was very feasible given - so the inpatient facilities have been doing this for more than a decade when you take this - having the reporting requirements, the latest to use the seclusion seems to be very feasible.

Peter Briss: Thank you.

Bonnie Zima: Yes. And I think again as discussed, this is a human recorded measure and we've already talked about, you know, the challenges that freestanding psychiatric units have in using electronic healthcare records.

Peter Briss: Sounds good. And anybody else have comments on feasibility? So let's move to usability and use please.

Bonnie Zima: Andrew?

Andrew Sperling: You want me to go first again? Yes. Usability, again (unintelligible) repeating myself, but a very usable, a very - and again, from other teams, what we mean by usability, it's usability for both the facility and for anyone who's doing comparative analysis and looking at quality of various facilities. The criteria like they're usable measure and a proxy for the quality parent side of the facility.

Bonnie Zima: Yes. And this is Bonnie. I don't have any unique comments that haven't discussed in the prior measure.

Peter Briss: Okay. Thank you. And anybody else have comments on usability in use? And how about - would anybody want to talk about related in computing please?

Bonnie Zima: I have no comment.

Andrew Sperling: No comment from me either.

Peter Briss: All right. Anybody have last comments that they'd like to make?

Jeff Susman: Hi. This is Jeff Susman. I guess the one thing that I'm still trying to square with both these measures is, is there room for further improvement? It seems to me the issues of seclusion and the state use of permeated healthcare, and that was made during meaningful changes. I'm trying to get just some

evidence or data that suggest yes, we still need these, other than the argument that there might be recidivism.

Peter Briss: Yes. So we did talk a little about - on the last measure in particular, whether things might be popped out. and one of the things that made me personally left were that that measure was topped out with the number - you know, the numerator of episodes of restraint use was still quite nontrivial number.

Jeff Susman: Yes. The question there of course is, what's the denominator?

Peter Briss: Right.

Jeff Susman: What's the right use?

Peter Briss: That's right.

Michael Abrams: Yes. So Michael here at NQF. I'm going to make one suggestion about the point that was just made by Jeff, you know, whether or not, you know, how you evaluate these, whether you think they're still usable. And it relates to the construct of harmonization and related in competing.

And it comes out of the suggestion that of course the two measures we just talked about are certainly related in some way. And actually kind of an interesting unique way perhaps, are they substitutes? Are they complements? Do you use them together to - as therapeutic interventions?

And moreover, when you add PRN meds to the mix and you're talking about overall chemical or physical restraints or seclusions, if when you are reviewing these measures, as you're getting ready to vote, and you think about usability, and you're concerned that just one of the measures may be leading

to increased rates of another measure like PRN meds, that's a negative effect that might lead you to suggest the measure is problematic and shouldn't continue.

Alternatively, you might want to keep the measure as is because it remains important as a singular construct. But perhaps in the future, and now we're thinking about sort of gaps in the harmonization, it would be nice to have a measure that brings all three together in some sort of composite form. So I just wanted to put those ideas out there.

Again, when you're voting, you want to think in terms of what you see on the page from the developer and what you know, in terms of whether you think it meets the various criteria. And in this case, I'm talking especially about related and harmonization and usability, okay?

Peter Briss: Yes. And we …

Michael Abrams: Fire away.

Peter Briss: Yes. And we - clearly different members of the committee had different set points on that. So that's going to need to be reflected in the voting, but I think we - for both of these measures, we've had a pretty robust discussion about them this year. And so - and we still have one more measure to try to do in the next 20 minutes. And so unless anybody has (unintelligible) on the last step, you better move us on.

Leslie Zun: It's Les, and I have a question. It's more kind of a procedural question in that even though my comments were fairly negative about the evidence in both restraint and seclusion, if our decision is just based on evidence, that is a

concern because I think these are good things to have. But if we're basing it on evidence, I question it.

And I question whether we're substituting medication or overmedicating versus treatment of this source, and we're not looking at any of those things as well, and we're not looking at chemical restraint. We're only looking at physical restraint.

So I think it's a good thing. It's just - I'm concerned about the evidence. And so procedurally, I would have to vote against it if I didn't think the evidence was there, even though I think it protects patients in the long term, is kind of the question.

Peter Briss: So I guess, Les, you're going to have to make a decision about whether you think the information presented gives you - gets you up to a level of moderate or high confidence that there's likely to be a net benefit of this measure as designed. And so only you can answer that question for yourself.

Leslie Zun: It's another dilemma. Thank you.

Peter Briss: That's what we do here. We assign dilemmas. Anybody else? I'd like to move us on. Does anybody else have urgent, burning things that they want to say before we move on? Hearing none, let's move to 1922. I have Dodi and Mady as the primary discussants on this one.

Dolores Kelleher: Right. So we need to hear from the developer first, Peter.

Peter Briss: Thank you. Developer first, Joint Commission.

Elvira Ryan: Hello. Elvira Ryan again. This is our HBIPS-1 measure screening. On admission, this has been implemented since 2011, NQF endorsed since 2014. This measure applies to patients who are inpatient psychiatry patients and also have ICD-10 codes for behavioral health.

The patient population again is patients aged one year and older. This measure looks to see that the patient was screened within three days of admission on five basic criteria. The criteria are patient strength, psychological trauma history, substance use, violence risk to others and violence risk to self.

Within the specifications for the criteria are descriptors that explain the type of documentation that is required in order for the abstractors to say that the criteria for the measures have been met. It is used in accreditation for freestanding psych hospitals, optional for psych units within an acute care facility. Thank you.

Peter Briss: Thanks very much.

Michael Abrams: So Peter, Michael here at NQF. So we just want to let you know, we are going to adjourn at 1:00. So we'll probably only just get started on this measure. And then we respectfully request that everybody reconvene at 2:00. And the developer, Elvira, as well, if you could come back at 2:00. We're sorry we weren't able to get through - we probably won't get through your entire set of measures by the end of this first meeting.

Peter Briss: Yes, I agree. So although this is related to the last piece, I hope we'll pick up speed over time. So if Dodi and Mady can walk us through as much of this as they can.

Mady Chalk: Go ahead, Dodi.

Dolores Kelleher: Mady, do you want to go or do you want me to go?

Mady Chalk: I want you to go.

Dolores Kelleher: Okay.

Vanita Pindolia: I'm sorry. This is Vanita. I just had a quick question, and I'm sorry to interrupt like this. I have a conflict with the afternoon. Is there a way I can submit my vote before I start my meeting in the afternoon? And this is just for NQF to think about it. You can just email me so I don't disrupt the meeting.

Elisa Munthali: Thanks, Vanita. We'll reach out to you.

Peter Briss: So Dodi, would you like to tee this one up first on evidence please?

Dolores Kelleher: Sure. Again, this is a maintenance measure that was first reviewed in 2013, assert in 2014. The developer provided evidence in the systematic review and specific to the measure quality, quantity and consistency, and it was grated.

In this measure, they did have updated evidence from the prior review and approval, which included a logic model and the updated APA practice guidelines for adults from 2016. The logic model presented so that screening formed the basis of a treatment plan, which would decrease the chance of psychiatric relapse, which would improve treatment and medication (unintelligible) and reduce the cost of ongoing recovery.

The guidelines were related to each of the specific screening elements. And

with evaluation of patient strength directly linked to the recommendation statement. (Unintelligible) cited in the guidelines as core activities for any initial psychiatric evaluation.

There was grading and each of the statements was either graded 1C or 2C. C represents the strength of the supporting research evidence and C represents a low rating. There was comment that this might be the case because there was great difficulty in studying assessment approaches in RCTs as noted in facilities.

Michael Abrams: (Unintelligible). Oh, I'm sorry. Go ahead.

Peter Briss: Mady, anything to add?

Mady Chalk: Yes. I have a couple of questions about the evidence. While it's nice that there's a logic model that shows that the screening will form the basis of the treatment - of a treatment plan from patients, there - we don't know that. There's no evidence of that in this. So that's number one. And that the treatment plan will then lead to reduced costs and better outcomes, there's no evidence of that either.

While I'm a great proponent of screening, I'm not - something about this does not compute to me. And - well, we'll get to the performance gap later. But on the basis of evidence, I'm not convinced.

Peter Briss: All right. So there's some concern that the screening hasn't been very convincingly linked to outcomes. Other comments on - from the committee?

Dolores Kelleher: This is Dodi again. I think there was some - if I read this right, there was some concern when it first came through about the - how directly the evidence

equated to the measure focus. But then the correlation of those elements was sufficient.

I think for me, this isn't any problem in these kind of process measures, because they purport to - unless there's a lot of evidence already in the literature about the outcomes, this is - so I think what Mady is referring to, there's really not a measure that really looks at whether the outcome is as suggested.

Mady Chalk: I mean the part …

Dolores Kelleher: And I don't know how to get around that, other than to have more outcomes related measures. But I think we - there's - this is not purporting to be an outcome measure if the process measure that indirectly or directionally says that it's related. I think it's certainly in terms of evidence, important to measure the - that there is this kind screening going on and that it has the potential for high impact. So I think we have the same problem we had initially when it was (unintelligible) for this problem.

Peter Briss: Any additional comments from the committee?

Leslie Zun: Yes. This is Les. I'm concerned about the - it doesn't make - this measure doesn't make sense. If I'm worried about a patient's violence and substance use, I'm not waiting three days to find that out.

Mady Chalk: You bet.

Leslie Zun: And so there's a problem with this measure just off the top of my head. I just can't understand why within three days, we're going to do this. But in the

interim then, they could be withdrawing. They can be hurting staff. It just - I think it's very problematic.

Dolores Kelleher: This is Dodi. Could the developer speak to why they picked three days?

Elvira Ryan: The measure states that the assessments should be done within the first three days. It doesn't mean that they should wait for three days to complete the assessment. There are times when the patients come in and they could be in crisis and maybe they cannot participate in the screening.

There's also parameters within the measure that if the patient is unable to respond, or if the patient is uncooperative with response, reliable significant other could provide information related to the screening criteria.

Michael Abrams: Thank you.

Peter Briss: Go ahead. Other comments on evidence?

Harold Pincus: This is Harold. And the question I have is, there usually - screening is usually tied in some logic chain to outcomes. And what are the actual next steps in the logic chain? What is the outcome about?

Mady Chalk: Harold, that was my question initially. The link - according to the logic model, the link is to a treatment plan. And then the further link is from the treatment plans to in treatment outcomes, not after treatment, in treatment outcomes of all kinds of things. But we don't know whether it was linked to the treatment plan. There is no evidence in this measure as such.

Harold Pincus: Okay. Thanks.

Mady Chalk: Does developer want to say something? I don't know.

Peter Briss: Any other comments from the community? All right. I think we've heard some diversity of opinion on this. I wanted to just - it says it's nearly one o'clock stop time. Why don't we take - unless the staff wants to kind of man this, why don't we break here and reconvene at 2:00 for the - for to finish up this discussion and do the last two measures?

Elisa Munthali: Thanks, Peter. That sounds good. Before we get off, I just want to take a quick pause to see if we have any public or member comment. Okay. I'm not hearing any, and I'm not seeing any hands raised. So with that, we can break and we will reconvene at 2 o'clock to continue the discussion of performance gap for measuring 1922.

Peter Briss: Thanks everybody.

Michael Abrams: Yes, thank you. We're more than halfway done, so that's good.

Elisa Munthali: Bye-bye.

Jeff Susman: Bye now.

END