NATIONAL QUALITY FORUM

Moderator: Kim Patterson
February 5, 2020
12:00 p.m. ET

(Hannah): Hello, has anyone joined the call?

(Peter): Hey, this is (Peter)

(Hannah): Hi, (Peter) How are you? This is (Hannah) with NQF.

(Peter): Hi, (Hannah) I'm terrific, thanks. How are you?

(Hannah): Good. Thanks for joining.

(Peter): You're very welcome.

(Hannah): Good afternoon. Thank you for joining the call. We are going to give committee members just a minute or two to dial in. We see a few people are still connecting so we'll get started momentarily. Thank you.

(Dodi Kelleher): This is (Dodi Kelleher), just to – in case other people are having the same issue, I couldn't go in the way I had the last couple of meetings which was to

hit, you know, call me and join the web so I had to just do the Join the Web and then call in directly and that worked…

Man: That happened to me. (Unintelligible), yes.

(Jody Keller): So people may be, you know, thinking they're stuck but they need – they can sort of do the web one and then call in the toll free number.

(Hannah): Thank you, Jody, appreciate that. We will try to figure it out on our end as well. Maybe we can send a note to folks to dial in using the phone line as well to try to navigate that. Thank you.

(Peter): This is (Peter) It might also save folks time as we're waiting for people to dial in, make sure you're log into (poll) everywhere.

(Hannah): Good afternoon. We are going to get started in just a moment. We are aware that there are some technical difficulties using the audio – the audio portion of the webinar, so we would encourage you to use the webinar link to access the platform and the slides, you can still (chat) that way but we would encourage folks to dial in using their phone line and we are sending a note to our committee members to make sure – as well as our developers to make sure that they will do that. Please do let us know if you're having any other issues. And we hope to get started in just a minute.

(Harold Lucas): This is (Harold) I think I've done it so that's good.

(Nicolette Mihas): Okay, (Harold), thank you. Good afternoon. Welcome, everyone, to the Behavioral Health and Substance Use, this is our webinar. This is our third measure evolution meeting. We thank all for joining to continue our measure evaluation discussions. My name is (Nicolette Mihas), I am the Director in the

Quality Measurement Department and part of the Behavioral Health Project team. And I am joined in the room by the rest of my colleagues here at NQF.

(Harold) or (Peter), would you like to share any comments with the group before we get started?

((Crosstalk))

(Harold Lucas): No, we just had our, you know, a portion of our meeting a few days ago so I think we can proceed ahead.

(Peter): And this is (Peter) Welcome to everybody who's joining today. Thanks for all the attention to date. We've made a lot of progress and we still have some progress left to make so let's try to be efficient today.

(Nicolette Mihas): Great, thank you. So to quickly run through what we hope to cover today, I guess I'll start with just sharing where we left off last meeting. So during our last call we finished discussing 3539E, That measure was recommended for endorsement. We reviewed measure 3541. That measure was also recommended for endorsement; and we started our evaluation discussions of measure 3492. That measure had passed the evidence gap in reliability criteria. It was consensus not reached on validity and so we will be re-discussing validity on our post comment call but we do need to pick up where we left off with measure 3492 and that is starting with the feasibility of that measure.

Before we jump in to feasibility, we do plan to get through the rest of our measures today. To do that we will – we hope to move through what we need to talk about efficiently. We will start by finishing discussions for 3492, then we will move to measure 3175. That measure is being reviewed as an ad hoc

evaluation and so we will only be looking at the reliability and validity. And then we will move onto measure 2800 and 2801 and those are two maintenance measures.

I think with that I would like to turn it over to (Hannah) to take roll.

(Hannah): Thanks, Nicolette. (Unintelligible) our co-chair (Peter) Britt.

(Peter): Here.

(Hannah): (Harold Lucas)

(Harold Lucas): Here.

(Hannah): (Mady Chalk) (David Einzig)

(David Einzig): Here.

(Hannah): (Julie Goldstein Gremit) (Carlos Gross) (Lisa Jensen) (Dodi Keller)

(Jody Keller): Here.

(Hannah): (Craig Nutson)

(Craig Nutson): Here.

(Hannah): (Michael Artieri) (Karen Marks)

(Karen Marks): I'm here.

(Hannah): (Raquel Jeffers)

(Raquel Jeffers): I'm here.

(Hannah): (Bernadette Melnick)

(Bernadette Melnick): Here.

(Hannah): (Lawrence Miller)

(Lawrence Miller): I'm here but I did send an email that I could only be here for 45 minutes today.

(Hannah): Thanks, Larry, we did get the email. Thank you.

(Lawrence Miller): Thank you.

(Hannah): (Brooke Parish) (David Pating) (Vanisha Panovia)

(Vanisha Panovia): Here.

(Hannah): (Unintelligible) Andrew Sparling.

(Andrew Sparling): I'm here, hello.

(Hannah): Hi. (Jeff Stutsman)

(Jeff Stutsman): Present.

(Hannah): (Michael Trangle) (Bonnie Vema)

(Bonnie Beamer): I'm here.

(Hannah): (Leslie Zune)

(Leslie Zune): I'm here. I'm giving a lecture in 45 minutes so I'll be on as long as I can. Thank you.

(Hannah): Thanks, Leslie. Did anyone else join the call who I missed or maybe joined during roll call? All right thank you. I ask that before you leave if you have to leave early please let us know either through the web platform via chat or you can go ahead and announce it verbally as well, thank you.

(Nicolette Mihas): Okay, (Mady Chalk), have you joined? We see your name in the platform but we didn't hear your name during the attendance. Just a reminder, if anyone is having problems with the audio please do reach out to our team either through the chat or via email. Thank you.

(Jody Keller): And again, if you go back into the email that was sent there is a toll free number to call in with an access code.

(Nicolette Mihas): Yes, thank you. And we just sent an email reminder about that so it should be at the top of your inboxes.

(Peter): And again, if you're not already logged into (poll) everywhere it might be a good time to do that now.

((Crosstalk))

(Nicolette Mihas): Okay, (Peter), I think at this time we will turn it over to you to start us off with feasibility for measure 3492.

((Crosstalk))

(David Einzig): …for the voting poll please?

(Nicolette Mihas): Can you repeat that?

(David Einzig): Oh this is (David Einzig), can you send me another link for the voting poll please? I need the link to log on to it.

(Hannah): Yes, we will do that, David.

(Peter): We are picking up where we left off on 3492, acute care due to opioid overdose. We've had a lot of – this is kind of a complicated measure and so we've had a lot of discussion already because we had consensus not reached on validity, we're already scheduled to have – to re-litigate this measure in the upcoming post-comment meeting, and maybe by that time we'll have the benefit of additional comments from other stakeholders, so I'd like – today I'd like not to re-litigate issues that we've already talked about, and we're just trying to finish up feasibility and usability and use, so if Jeff or (Tammy) could walk us through feasibility please?

(Jeff Stutsman): This is Jeff. I can do it, (Tammy), if you want.

(Tammy): Great, thanks Jeff.

(Jeff Stutsman): Because I think it's pretty straightforward, I do think this is a fairly feasible measure that's part of a routine care if you're admitted to an ER and the linkage with codes would suggest opioid overdose I think are clear enough.

(Peter): (Tammy), anything you'd like to add?

(Tammy): Yes, the measure used is Medicare claims data. There was one comment that (unintelligible) but it's using Medicare claims data.

(Peter): Anybody else have issues or concerns that we ought to talk about before we vote?

Leslie Zune: This is Les Zune. My concern is – and maybe this was addressed earlier, but when patients present to emergency departments with opioid overdose as the criteria noted here, frequently they do not put that as the final diagnosis for multiple reasons. And so I think that it's important that we understand that we may be missing a lot of patients because, you know, if they fell down and broke their ankle, if they had a respiratory arrest, if they had something else that's going to be the diagnosis for a number of reasons, one which is insurance companies may or may not pay if one does that or not.

(Tammy): Is that true for primary and secondary, because this measure uses any diagnosis, not just the primary.

(Larry Zune): I can't say that for sure.

(Danita): This is (Danita), I can tell you we did a – I did a whole search for two years to try to figure out how many opioid related hospitalizations or ED visits we have using the codes that are available using as a primary discharge and then just looking at any point as a discharge diagnosis code and we got such a

small number of patients that it looks like we have no opioid issue at all which isn't the case, it's just they're not coding it accordingly.

(Tammy): Yes, I think they actually say in the validity – actually this is part of the reliability analysis that they did a electronic medical record review and found that the sensitivity was very good, which is what you're saying that if someone has an overdose, you know, but they're not coding people as an overdose who don't have it. But that the sensitivity is not…

((Crosstalk))

(Tammy): …a lot, that's basically.

((Crosstalk))

(Tammy): …the conservative estimate. But the question…

((Crosstalk))

(Tammy): Yes, but it's a reliability issue, I think not a feasibility issue.

(Peter): This is (Peter) I sort of agree that this issue is an important issue. I would have put it into reliability and validity which we're already going to need to re-litigate at a later call so I'd actually like to table this specific issue and – and for when we talk again about validity and talk to us about feasibility issues here.

Raquel Jeffers: So this is Raquel. I can see in the pre-evaluation comments that someone raised that 42CFR Part 2 as an issue, but I don't understand if a hospital is reporting aggregate numbers of overdose steps from a population sampling

why 42CFR is a barrier? I'm just wondering if someone could address their concerns?

(Meti Tak): This is (Meti) I don't think that's the issue – an issue.

Raquel Jeffers: Okay.

(Meti Tak): The way I understand 42CFR.

Raquel Jeffers: I agree. Okay.

(Peter): Anybody else have issues on feasibility? Why don't we try to move to a vote please?

(Hannah): Thank you. Voting is now open for feasibility on measure 3492. Options are A for high; B for moderate; C for low; and D for insufficient. I'm seeing 15 votes; I think we're waiting on one or two more. Is anyone having issues with the voting links? Voting is now closed for feasibility on measure 3492. We have 2 votes for high; 11 for moderate; 2 for low and zero for insufficient. This measure passes on feasibility.

(Peter): And that takes us to usability and use. (Unintelligible)

(Jeff Stutsman): You want to take usability and use separately or together?

(Peter): We'll have…

((Crosstalk))

(Mike Levy): (Mike Levy), I just joined the call. Sorry about that.

(Hannah): Thanks, (Mike)

(Peter): Thanks, (Mike)

((Crosstalk))

(Peter): …have done it already – (Mike), if you haven't done it already you should pull up the poll everywhere.

(Mike Levy): Yes, I'm just pulling it up now.

(Peter): Jeff, I think if you're inclined we could – we're going to have vote usability and use separately. I think it's fine to sort of discuss them together.

(Jeff Stutsman): Yes, I mean, this measure has been reported, the Maryland experience, for example. If – I don't know if you'd really call it an accountability application but I'm satisfied that this could be used for improvement. We've already talked about the many challenges to comparisons so I'm not going to reiterate that. I think if you strongly feel that's a problem then you might be concerned that there's a potential negative consequences that a measure not, not reflect reality without a number of caveats. But again, we've already done that.

And that there is some I think clear linkage that could be made toward achieving improvement in care, you know, and that's really I think – it should, in my mind, pass on these. I'm not – I don't think they're high – high assurances but I think it's okay. (Tammy), do you want to…

(Peter): (Tammy)

(Tammy): No, just point out someone wrote in SAMHSA requires states to collect and submit emergency room opioid overdose data for the purposes of allocating federal funds. All states have been doing this since 2016.

(Peter): Anybody else have comments on these dimensions? If not let's try moving to voting please.

(Hannah): Voting is now open for use on measure 3492. Options are A for pass; B for no pass. We're just waiting on one more vote. Voting is now closed for use on measure 3492. We have 13 votes for pass; 3 votes for no pass. This measure passes on use.

(Peter): And unless anybody wants to make an additional argument I think we can move right to voting on usability. I mean, yes, usability.

(Hannah): All right, voting is now open for usability on measure 3492. We have options are A for high; B for moderate; C for low; and D for insufficient. Voting is now closed for usability on measure 3492. We have 2 votes for high; 10 votes for moderate; 4 votes for low; and 0 for insufficient. This measure passes on usability.

(Peter): And again since we – since we didn't reach consensus on validity we – I think we don't vote on an overall vote at this time while we wait for other folks to comment and so does that finish our discussion for today on this measure?

(Nicolette Mihas): Yes, correct, (Peter) Thank you. This is (Nicolette) from the NQF staff. I did just want note for the record that measure 3492 did receive three public and members comments before this call. Those three comments were supportive of the measure. You are correct, we will not be voting on overall suitability for endorsement at this time but the measure will move forward to the

comment period and we will revisit the validity criteria on the post comment call. So we can move onto our review of measure 3175.

(Peter): And I'll pass the baton over to (Harold)

(Harold Lucas): Okay, now I understand this is an ad hoc review and maybe – just to explain the context of that.

(Hannah): Sure. So for measure 3175 as noted in the measure worksheet as well as on the orientation call, this measure is brought forward for an ad hoc review. Our policy for ad hoc review is that it's a formal measure evaluation outside of the scheduled maintenance period and so an ad hoc review is limited in scope and focused on a specific issue regarding a certain evaluation criterion. In this case we will be looking at the scientific acceptability so the reliability and validity of measure 3175 at the individual clinician and the clinician group levels of analysis. This is one of the material changes of a measure that can trigger an ad hoc review.

We also did want to note, as background, that this measure did go to the measure application partnership in 2018 for potential inclusion in the (MIPS) program. The recommendation from the (MAP) at that time was for the measure to be refined and resubmitted, they directed that before implementation in (MIPS) that the measure was tested and evaluated by the relevant standing committee of NQF specifically for reliability, validity as well as the attribution methods. And so that is additional background as to why we are looking at those criteria for today.

(Harold Lucas): Okay, can we hear from the measure developer with specific attention to those, you know, key issues that were raised?

(Soren Matthew): Okay, hi. This is (Soren Matthew) from USC, can you hear me?

(Harold Lucas): Yes.

(Soren Matthew): Good. So thanks to Nicolette for explaining this unusual request, I have to apologize for this narrow scope of review. As she said, we are submitting this material in response to what the (MAP) asked us to do and what CMS agreed to fund and is only really only looking at the measurement properties at the clinician and clinician group level.

Conceptually I don't think I have to lecture this panel about the severity of the opioid epidemic and the importance of medication as (unintelligible) treatment as an evidence-based but underused option. However, as patients who are starting on (MAP) often discontinue treatment prematurely, we have developed this continuity measure for it two years ago. And the measure captures whether patients remain on pharmacotherapy for at least 180 days with no gaps of greater than seven days.

The choice of 180 days as the duration was made because most drugs that were tested in the FDA approval process for that period so we do not have evidence of effective usage for the durations and we had chosen a minimal – a maximum seven day gap requirement because of evidence for excess mortality with treatment gaps in that range of about one week. And this is because patients were desensitized to opioids during pharmacotherapy are at elevated risk of overdoses when they relapse.

Unfortunately we continue to see, as you saw in our submission materials, pass rates below 30% so we are well away from good performance here. And not to be specific elements that we submitted with the additional testing. So for attribution we are in the fortunate situation that this is a pure prescribing

measure so all the actions in the numerator are (unintelligible) administration or prescribing of pharmacotherapy and so we can directly identify the prescriber from the claims and then attribute the patient based on the prescriber.

As it turns out, the majority of patients have only one prescriber for their pharmacotherapy and about 2/3 of patients get all their prescriptions from the same practice, so attribution actually is in our view pretty straightforward.

We conducted reliability testing with standard Rand method, and thought reliability sufficient for both clinician and clinician groups. And we conducted validity testing using an expert panel approach with supportive results. We do acknowledge the point that many clinicians had small denominators but want to qualify that the used Medicare data – Medicare fee for service data for testing and therefore only see a subset of each clinician's panel.

We expect to have many more clinicians qualifying for the measures implemented in (MIPS) where a registry would be used for all payer claims would be used. And that's particularly true because addiction care still tends to be in the hands of specialists. There's a lot less variability in the types of patients and indications that are seen in specialty practice so therefore patient numbers with the same diagnoses and treatment tend to be larger per panel which makes it easier to have denominator sizes that are sufficient to construct data measure rates.

I'm looking forward to today's discussion.

(Harold Lucas): So I believe (Julie Grommet) and (Dave Panning) are the lead discussants. Do you want to kick US off, here?

(Dave Panning): Who do you want to start? I guess, you want Jeff or me? Either one of us?

(Harold Lucas): Sure.

(Dave Panning): Okay. This is David Panning. Looking at measure 3175. This measure looks at the percentage of adults at least 18 years of age who have been on pharmacotherapy for opioid use disorders who have 180 days continuous treatment. So we're going to look at just reliability and validity.

In terms of this construct of this measure it's a look back of claims data in Medicare so it's a fairly reliable data source. In terms of the exclusions, they look at – that there is no more than a seven day gap in treatment supply, and this would be a buprenorphine, naltrexone, buprenophine and naltrexone, doesn't include the extended release ones and methadone.

So, what else was there? The typical denominators were fairly standard ICD-9 codes for opiate dependence and opiate related disorders. So the conclusion was that the data elements are pretty clearly defined. The comments that a group gave was that the – these data elements were collected mostly on clients that were over 65. There were some aspects of the demographics and the male to female ratio which was one to one, makes you wonder about what – this may be more of a validity issue whether it's measuring through opiate dependence given that the ratio of men to women is usually like two to one or even three to one.

There were some concerns about the applicability for this to rural areas and some concern that health plans may not have tax IDs so – but this was actually not a health plan measure, this was a clinic and group practice measure. So with regards to our reliability specification 2A-1, there looks to be very consistent data collection with very clearly defined exclusions.

So I'll stop there. Are there any questions regarding 2A-1, reliability specifications?

(Tammy): I had a question about the tax ID. My understanding is this is a claims measure and the way that they attribute the provider to the clinic is through the tax ID and I don't think that's commonly available, so I was just wondering if the developer could speak to that?

(Soren Matthew): Yes, this is basically how the (MIPS) system does it. They have the tax ID under which individual clinicians report and they use the tax ID to which clinicians report the majority of their claims in Part B, so that's basically following their method of attribution.

(Mike Levy): Yes, this is (Mike Levy) Tax IDs might get you mixed up if it's submitting through a group. Do they look at the (NTI) number because that gets more specific to the individual provider.

((Crosstalk))

(Soren Matthew): We do do the specific provider through the (MPI) and then we roll up individual providers by tax ID to clinics and then we report on both levels.

(Mike Levy): I got you, thank you. Yes, that may help clear it up.

(Harold Lucas): This is (Harold) I had a question about how did – you said that 2/3 have a single provider but what did they do about the 1/3 who where there's more than one provider that's doing the prescribing?

(Soren Matthew): You mean in terms of attribution?

(Harold Lucas): In terms of attribution, yes.

(Soren Matthew): Yes, so what we do we do pluralities of days covered, so the provider or clinic who has the plurality of the days under treatment is the one who will receive the patients in the attribution method. And that mirrors the way that it is typically done based on visit patterns because the patient can see more than one doctor and then if you attribute based on visit patterns you basically attribute the patient to the clinician whom he or she sees most.

(Dave Panning): Yes, I was actually very critical of that when I first read it, and then I went into your attribution report which was very well written both for the group level and individual level attribution so you kind of have to look at that additional data that was supplemented to answer that question, (Harold)

(Harold Lucas): Okay. Other comments about reliability?

Woman: I just have a question, can you clarify whether you're looking your definition of outpatient includes primary care?

(Soren Matthew): Yes, any encounter.

Woman: Only outpatient encounters. And then…

(Harold Lucas): If people could mute their lines when they're not speaking because there's some background noise. Other questions about reliability?

(Dave Panning): Well we're just talking about specifications, not the testing. Did you want to review the testing? So let me go into reliability testing 2A-2. The provider level had (unintelligible) reliability of .77 and the group practice had a

reliability of .76 so passing the general threshold of .7 reliability. It's interesting because the state reliabilities were really, really high, they were like .95 or I can't remember the exact number. So this is a decrease but there was still high reliability at this level.

I think one of the issues that came up however, was based on the specifications, there are large exclusions so some 37% of the group practices were excluded and 48% of the clinicians were excluded, so – maybe I can ask the developer to speak to that issue and how does that impact the reliability?

(Soren Matthew): Well you cannot calculate reliability if there's no more than one patient and therefore we couldn't include a subset of the – of the sample in that analysis. As I said on the outset, we are using only one payer; we are using Medicare fee for service data and of course Medicare fee for service only represents a subset of the patients in a physician's practice or in a clinic. So we assume that with an all payer claims or with a registry data based construction we would be able to include more patients.

Reliability numbers tend not to be that sensitive to the sample size like I wouldn't expect that by including more clinicians reliability would come down, but obviously that is an empirical question.

(Harold Lucas): Is there a minimum number of patients that need to be – for a provider to be eligible for this?

(Soren Matthew): Mathematically it needs to be at least two patients because otherwise you cannot construct any variance. We tried minimum numbers, we tried just two everybody, we tried 25, which is the reporting threshold, and the reliability scores did not really change much.

Woman: If I understand correctly then so Medicare patients is a subset of the substance use disorder population and half of – almost half – 48% of the practices have less than two patients. So the reliability testing is done on a small subset of a small subset of the substance use disorder population. Why didn't you use Medicaid as your data set?

(Soren Matthew): So first of all we would calculate reliability for the scores that would be reported, so I think that's a valid approach because if a physician has only two patients his or her numbers would not be reported for concerns about stability of the rates. The choice of data of course needs to be pragmatic for independent developers like ours because we cannot just access data as we wish. And in this case we had to work with Medicare data.

(Dave Panning): (Harold), so these questions about the subset of the subset, I just ask for maybe guidance from the group, like the male to female ratio is one to one which is not typical of drug using populations; the 48% exclusions do make you wonder are they picking a subset of that subset? Are these Medicare – are these methadone clinic clients or what kind of practices do they have more than two or four over 65 clients on – in maintenance programs?

And then the Medicare aspect of majority of patients are over 65, are these validity questions that we need to be address next or more feasibility and usability questions because of the small subset?

(Harold Lucas): So I think it has to do with both. And I guess the question as I would frame it is does the population on which these – this measure was tested, does it represent the population to which it would be applied? Or…

((Crosstalk))

(Soren Matthew): Let me qualify one thing because there's a lot of talk about patients over 65. The vast majority of our patients are not 65 and older, they're below that age group and they're dually eligible. So there's about 50,000 and change in the denominator or 60,000 – 70,000 in the last measurement period and 60,000 out of 70,000 are actually younger than 65 and those are typically either disabled or dually eligible.

(Harold Lucas): So, but, (Soren), could you sort of answer my question in terms of, you know, what – how does the population to which this measure would be applied differ from the population to which would be tested?

((Crosstalk))

(Soren Matthew): And again, if we look at the type of people whom we have, which is predominantly patients who gained Medicare eligibility via a disability as opposed to old age, and predominantly folks that are dually eligible, i.e. have Medicare and Medicaid, I think the vast majority in our sample represents your typical opioid use population as opposed to your elderly Medicare population whom OUD is much less common.

(Harold Lucas): I mean, is there data to make that estimate? Because, I mean, there's a lot of individuals who are abusing opioids who were in, you know, particularly in the younger subset who may not qualify for disability.

Woman: I think it would be fair to say that the population that the measure was tested on is more disabled than the population the measure might be applied to following how you framed the question, which I like.

(Soren Matthew): I mean, keep in mind that we tested the measure in both Medicaid and commercial for the initial endorsement and we basically see the same patterns

there. This is really just using Medicare data to do individual clinician and practice attribution and test whether that's feasible and generates reliable numbers.

(Dave Panning): This is David Panning again. Could I ask for the group practice cohort were many of those methadone clinics or opiate treatment programs that had over 25?

(Soren Matthew): We can't tell; there's no designation of such in the claims pattern. I would assume that – I mean, methadone is not that commonly represented so my assumption is that there's going to be a lot of wayward practices that represent buprenorphine which is sort of the vast majority of the patients like about 50,000 out of 70,000 are on buprenorphine and with our without naltrexone.

(Tammy): This is (Tammy) Marks. My understanding is that Medicare didn't cover methadone clinics until I believe this year they started adding it so they're probably not in the claims data at all, but they will be.

(Dave Panning): Oh, thank you. All right, so…

((Crosstalk))

(Dave Panning): Yes.

(Harold Lucas): Any further discussion about reliability? So I think we're ready to vote. Are we setting up the vote?

(Hannah): Voting is now open for reliability on measure 3175. Options are A for high; B for moderate; C for low; and D for insufficient.

(Tammy): Can you just remind us what passes and what doesn't?

((Crosstalk))

(Hannah): Sorry…

((Crosstalk))

(Hannah): Votes for high or moderate would pass.

((Crosstalk))

(Tammy): Thank you.

(Hannah): Voting is now closed for reliability on measure 3175. We have 2 votes for high; 11 votes for moderate; 3 votes for low; and 0 votes for insufficient. This measure passes on reliability.

(Harold Lucas): So now discussion on validity.

(Dave Panning): Hi, with regards to validity, the developer relied mostly on (phase) validity. I believe they had eight – no nine members on their panel. One-third of the panel did not agree with the validity measures and there wasn't really explanation. I guess the big question was could they explain of the nine results of (phase) validity five strongly – no, one strongly agreed, seven agreed and two neither agreed nor disagreed. So there was no disagreement with the (phase) validity but there was concerns at the level (unintelligible) This is provider level I believe it is about the (phase) validity.

So maybe I would ask the developer to comment on that because there was a strong request to hear why those folks did not feel the measures were valid at the provider – oh at the – yes, provider and group practice level.

(Soren Matthew): I mean, to emphasize the vast majority of people, almost 2/3 agreed with validity or strongly agreed. Then there were a few ambiguous responses. We didn't get any specific comments as to why they voted, I mean, we got a lot of comments on why people voted positively but we unfortunately did not get anything on why they voted negatively or ambiguous.

(Dave Panning): I guess my concern is given that you have nine people, that's the only evidence of your validity and 1/3 of the nine neither could agree nor disagree nor comment on this measure makes me wonder about the validity of it as well as maybe why you also did not submit any – you have an enormous data set I would imagine and there might have been some way to develop some other measures of validity that you could have provided us other than the kind of lackluster (phase) validity which was reported.

(Soren Matthew): With the (phase) validity method is the standard method in that phase of the (unintelligible) endorsement so correlation and so construct validity testing is not part of this phase of the evaluation.

And I want to emphasize we are not talking about the validity of the measure, we are talking about the validity of using the measure for provider/clinician level measurement. The overall validity had been endorsed previously.

(Nicolette Mihas): This is Nicolette from NQF. I just wanted to add that for – when a measure is first coming forward we do accept (phase) validity results as a method of testing that is acceptable. I did want to note that for this measure it will be due

in 2020 or 2021 for maintenance review at that time empirical validity testing will be required.

(Soren Matthew): Yes, and we plan accordingly.

(Harold Lucas): But, (Soren), you said that there was validity data presented in its original submission?

(Soren Matthew): Well not construct validity because back then this wasn't an NQF requirement, this construct validity testing is a new thing. But back then the measure passed on validity again using an expert panel approach.

(Harold Lucas): I see, okay, just wanted to clarify that. Any further discussion of validity?

Man: I'd just like to ask regarding this process, so assuming we pass this measure either on reliability and validity or both, this goes then to further – this really is sort of a path to further development in which time you would submit more – I guess this is construct validity but you would give us more – more – I'm not sure what they call the other (unintelligible) validity but more data with regards to validity testing, is that correct?

(Soren Matthew): Yes, it's a bit tricky. I mean, again, us independent developers have to sometimes go with what data and what funding we can get. So for 2020 we are up for a full maintenance review and we have planned and budgeted for construct validity that would use our rates and correlate those rates with other measures to show that the construct is actually valid rather than just expert panel rating.

(Sam): And hi, this is (Sam) (unintelligible) with NQF, just a brief follow up related to that question. So this measure was brought to NQF specifically to evaluate

its feasibility for inclusion into (MIPS) so our task today is to take a look at this based on what the measure developer has provided, but ultimately this will be due for maintenance either in a cycle or two where the measure developer will be required to submit more substantive validity.

For maintenance of endorsement we require empirical validity testing so that needs to occur either at the data element or score level. But for this review the (phase) validity will be sufficient.

(Jeff Stutsman): I have a concern – it's (Jeff Stutsman) – that the move to (MIPS) will be based on (phase) validity of a very small number of experts with some ambivalence among them. And that even though this is going to undergo a maintenance review relatively soon, I would assume, that this would be implemented prior to such further construct validity or further testing. I just worry, I think that when we move to accountability at an individual or clinical organization level we should have perhaps more rigorous assurances that we have a valid measure. And I appreciate the developer's careful and articulate description of what's been done but I do worry some.

((Crosstalk))

(Soren Matthew): It's not a requirement, we didn't do it so this is a little tricky for us to be asked to do something that's not in the submission requirements.

((Crosstalk))

(Jeff Stutsman): Oh I understand.

(Harold Lucas): Yes, could I just ask if, NQF staff, because (unintelligible) we were to pass this measure, would it then open up the likelihood that CMS could put this

into (MIPS) for this year or for next year? And then the reassessment would occur after that it's been put into (MIPS) Is that a possible pathway?

(Sam): Sorry, could you reiterate that? I didn't quite catch everything.

(Harold Lucas): So the question is if it's passed today by our voting on reliability and validity, is then – does that essentially go along with the recommendation of (MAP) that at that point that it passes our vote, they would be fully supportive of its implementation into (MIPS) this year?

(Sam): You're correct.

(Peter): Well except – this is (Peter) There are more steps than in that process generally speaking than were reflected in that answer. So the (MAP) has looked at it at least once, they have said, you know, contingent on sort of (phase) validity that this is measuring something worth measuring at the provider or group level, they thought that the measure was generally worth the – sort of considering for rule making. CMS would then consider it for rule making if it passes muster at CMS they would generally sort of put it out as – in the federal register as measures under consideration and based on public comments they would make a decision.

So the short – that was a long answer. The short summary of that is that there are several more hoops that it would have to go through for CMS to include it in the rule making.

(Harold Lucas): Yes, I'm just trying to figure out if those hoops would sort of extend the timeline so that it would be overtaken by their resubmission with more data.

(Peter): Yes. I mean, to this – this one seems like – if the question is – the main question it seems to me is sort of can this kind of a measure be attributed at the individual clinician level and truth is this one is – seems simpler to me to attribute to a clinician than many other measures that NQF endorses. So it's a – sort of – although this has been a really good discussion, and it seems to me that this one's not such a hard one to attribute.

(Harold Lucas): Any other comments? Okay, so I think we're ready to vote on validity.

(Hannah): Voting is now open for validity on measure 3175. Options are B for moderate; C for low; and D for insufficient. We need just one more vote.

Man: Remember that I can't vote just in case you're looking for my vote.

(Hannah): Thank you. Thank you.

((Crosstalk))

(Hannah): So based on the voting results we only have 14 committee members able to vote at this time. Our quorum number if 15 committee members, so we will not be able to share the results of the vote at this time. We will have to send out an electronic voting survey. We can send that right now to members on the call, we will also send it to the other committee members along with the recording for this meeting in order to make sure that we receive the required number – quorum number of at least 66% of our committee members present to be able – for the vote to count.

Woman: So if we were on the call and voted, then we don't have to follow up with the electronic request and it's just people who are not present right now?

(Hannah): So we would ask that you do vote. We are going to send out the voting survey as, I believe it's a Survey Monkey link, at this time. We would ask that you do recast your vote using that platform so that we have it captured. We will not, from this point forward, we will not be able to share any of the results of the vote until we make sure that we have all – until we reach that quorum number. And so those on the call will be able to vote using the link and we will also share those – share the link with those that are not present in order for them to review the meeting recording and transcript and cast their vote as well.

((Crosstalk))

Man: A lot of wasted time.

(Peter): All right so this is (Peter) again. I think that moves us to 2800. So we still got two more – two more maintenance measures to go in an hour so I think that's a, you know, or in a little bit less than an hour so I think it's a doable thing. We'll have to continue to work on being efficient. So with that – would the developer like to tee up 2800 for us please?

Man: Can I just a point of clarification?

(Peter): Sure.

Man: So all future votes now have to be followed up with the online polling?

(Hannah): That's correct.

Man: In other words – okay.

(Peter): Yes, the short answer is yes.

Man: Yes, okay. Thanks.

(Peter): So with that, would the developer like to tee up 2800 for us please?

(Emily Morden): Hello, this is (Emily Morden) from NCQA, can you all hear me?

(Peter): yes.

(Emily Morden): Great. Okay, so just going to start off by kind of highlighting for both of the next measures that we're going to be discussing, that we know that the safe and judicious use of antipsychotic medications is really a critical issue for children and youth. And we know that the – that antipsychotics are a powerful medication that are really only indicated for treating a limited range of mental health conditions in children and adolescents. And we know they also have the potential for serious side effects.

So to address these concerns we've developed the two measures we'll discuss today. These measures really encourage an approach to think before you prescribe and then when you are prescribing to carefully monitor youth who are on antipsychotics.

So this first measure, 2800, metabolic monitoring for children and adolescents on antipsychotics, it assesses the percentage of youth with ongoing antipsychotic use who received a blood glucose and a cholesterol test during the year to monitor for the metabolic impacts.

The health outcome risks of antipsychotic medications are well documented and metabolic monitoring is really an important part of managing those risks and it’s recommended by multiple clinical practice guidelines.

The measure is intended for use in roles in Medicaid and commercial health plans and it has been used in (HETIS) for health plan reporting since 2015. And beginning this year the measure will also be included in the Medicaid Child Core Set or state reporting as well. So thank you.

(Peter): And maybe one more follow up question, this is (Peter) Since this is a maintenance measure are there – are there areas in this measure where you believe that things have importantly changed that we should be aware of since the last time the committee looked at this measure in '16?

(Emily Morden): Yes. Just a couple things to note, since the last time it was reviewed for endorsement, we have made a change to split out separate rates that are reported to look at the two specific metabolic tests, so a separate rate looking at receipt of glucose tests versus the cholesterol test and then also we still have a total rate but since the measure has been out there in use this was an area that was identified that's helpful to be able to target quality improvement efforts and understand where the gap in care is particularly for children here.

And then the second change that I'll note was relatively minor but we did combine two of the age strata in the measure so now there are just two strata, one of the ages 1 to 11 and then a second one that looks at children ages 12 to 17. And then of course we have a total rate as well.

And we really combined those due to the small number of children that we saw in that – the lowest age strata and so it really improved the report-ability when we combined it with the 6 to 11 year olds.

(Peter): Thanks very much. That's very helpful. So I have the – the lead discussants as (Bernadette Melnick) or (Bonnie Vema), if you guys would like to move us

forward on evidence, that'd be great. And this can be – unless you feel like something is particularly concerning or something may have changed since the committee last considered this measure, we might be able to spend a little less time on these criteria and a little bit more time on usability and use.

((Crosstalk))

(Bernadette Melnick): Yes, (Bonnie), do you want to proceed or would you like to (unintelligible)?

(Bonnie Vema): (Bernadette), you can go first.

(Bernadette Melnick): Okay, fine. There weren't really any major changes to the evidence. And the committee overall was supportive of the importance to continue measuring this measure.

(Bonnie Vema): And this is (Bonnie) Just to clarify, the measure description is glucose and cholesterol, not or?

(Bernadette Melnick): I thought it was both.

(Emily Morden): Yes, this is (Emily) I can help clarify. So we actually looked for both in the measure. We do have separate rates so we looked to see if kids receive glucose tests, then we also looked to see if they received cholesterol. And then the total rate looks to see if they received both recommended tests.

(Bonnie Vema): Right, so but the main measure is an "and" is that right?

(Emily Morden): Yes. We're really looking for receipt of both of these recommended tests.

(Bonnie Vema): And in the core set it went in as an "and" or an "or"?

(Emily Morden): The core set specification will be aligned with the most recent, you know, (CITA) specification that we provided as part of the materials, so it will be looking for those same three things.

(Bonnie Vema): Is it an "and" on the core set?

(Emily Morden): For the total rates, yes.

(Bonnie Vema): So it would be both to pass, you have to have glucose and cholesterol to pass?

(Emily Morden): Correct.

(Bonnie Vema): Okay. The other thing too is I just wanted to bring to the developer's attention, I noticed in the submission that, you know, it was sort of a repeat of that 2012 treatment guidelines from the Academy of Child and Adolescent Psychiatry, and it said there was no new evidence. So I'm a little surprised because even a Google search, you know, yielded three more papers, (Delante) in 2014, (Gratu) in 2015, and a recent study by (Jensen) which might not have been possible to add because it just came out in November 2019.

But the bottom line is that from this new body of evidence it suggests that there's – I just need to share this – there's variation in risk for elevation of these parameters by type of antipsychotic. And I'm not going to bore people on the major findings, but encourage NCQA to at least do a Google search on this and beef up the lit reviews as 2012. So those are all my comments on evidence.

(Peter): Thanks, (Bonnie) Any other comments from the rest of the committee on evidence? Hearing none I think we might be able to move to a vote please? Oh or will we even vote or shall we hold the voting until the poll comes out?

(Hannah): Yes, so we will go ahead and vote like usual. We have sent the Survey Monkey link to the committee members that are on the line, so if you would like to use that to cast your vote as we're going along we do encourage you to do so. But we will not be announcing any voting results at this time since we do not have quorum.

(Peter): Oh and I'm sorry…

((Crosstalk))

(Peter): So if we vote – so if we vote in Survey Monkey now it sounds like we might not have to respond to Survey Monkey again later, is that correct?

(Hannah): That's correct.

(Meti Tak): Oh, okay that's what I was asking before, okay. Great.

Woman: So can we just vote in Survey Monkey now or you need us to do it in both places?

(Hannah): No, just in Survey Monkey now to cast your vote.

(Peter): It came out – if people are looking for it now…

((Crosstalk))

(Peter): …at one o'clock in my email.

(Maggie): Well, hasn't gotten to me yet. This is (Maggie), then.

(Harold Lucas): This is (Harold) I'm conflicted so I assume I would not be getting the Survey Monkey email?

((Crosstalk))

(Maggie): Here it is.

(Hannah): (Harold), we did send you the voting link for – because you're not recused on 3175 so you can cast your vote for that measure.

((Crosstalk))

(Harold Lucas): Oh okay.

(Hannah): And then for 2800 and 2801 there's an option to recues yourself from voting if you're conflicted.

(Harold Lucas): Okay, but I already voted on 3175.

(Hannah): Okay. Yes, so please do not cast your vote on 2800 or 2801, we will not be able to…

(Harold Lucas): Right.

(Hannah): …accept it even if you do so.

(Harold Lucas): Yes, so…

((Crosstalk))

(Peter): So let's – so either that Survey Monkey link is already in your inbox or should get there shortly, I think I'd like a – I think I'd like not to wait for more of that while we continue to the discussion if everybody is amenable to that. And if everyone is now clear on how we're going to vote.

Does anybody else have questions about process or voting procedure?

(Jeff Sussman): This is (Jeff Sussman) This means we're not going to use Poll Everywhere, but we're going to use that link and if we do that link now that will count as our final vote. Is that correct?

Woman 1: That is correct. We are not using Poll Everywhere we are using the Survey Monkey link.

For the last vote on 3175, we would ask that if you have casted your vote using Poll Everywhere that you do make sure that you cast your vote using Survey Monkey. We need it to be in the Survey Monkey format in order for us to count that vote.

So we would ask that you repeat that one in Survey Monkey. Moving forward we will only use the Survey Monkey link.

Man 1: Any other questions or concerns at this point?

((Crosstalk))

Man 2: Which question I'll be on that we're voting on?

Man 1: We just finished the discussion on the importance of measuring reports and we're going to move next to performance gap on measure 2800.

Man 2: Got you. Thank you.

Man 1: And if there's a fair amount of background noise so if you're if you're not currently speaking, I'd appreciate if you could mute your phone.

All right. So can we move to performance gap please? (Bernadette) or…

(Bernadette): Yeah, the performance gap has been demonstrated. So there is opportunity for improvement.

(Bonnie): And this is Bonnie just to let you clinically this, the measure the way it's constructed, it sets a very, very low bar for medication safety monitoring. Meaning that if you get a glucose level the cholesterol anytime during the year that a kid has two AP meds within a 90 day period.

It's not really aligning with what clinically we're trying to do and that is that a child is on ongoing in a psychotic medication and you check baseline glucose cholesterol, you check it again, among those with continuous care. So the performance is low and that is capturing even what I would say it isn't really monitoring.

Man 1: Just for the rest of the committee even with a relatively low bar, the performance rates are on the order of 33%.

(Bonnie): Thank you. Thank you.

(Venita): This is (Venita) I had one question for the current QA. So whenever we get the measures that are for renewal, I am always interested in understanding if we're making improvement or not, and as in this one as it's been stated it really isn't improving. And even though the bar is set so low – so is that why was there discussion to see if we can further improve by separating and giving providers the knowledge of what it is?

I would hate to just keep renewing something and really not understanding the root cause of what might be causing it. Is that what the – is that what was taken away? Is that maybe if we record it as a measures recorded as in hand, but if we gave the information then or then the providers would know what's missing? Is that the intent?

Woman 1: Thank you. Yes, that that is one of the reasons why we saw it important to split out these two tests. We've actually done some work on this measure with health plans and doing quality improvement, and what they found was that there's even probably the larger gap in care for getting cholesterol test performed.

And in, we don't, we aren't able to see that unless we actually separate that out as its own performance rate. So that's one of the reasons why we did that.

And then the other thing I'll add is that we've also kind of looked at this measure and looked at, you know, regional variation, and also tried to identify from our (CEDUS) data, you know, plans that are performing well, and what they might be doing.

We have seen an association for this measure, in particular, that health plans that operate in states that have prior authorization programs perform consistently better, or significantly better, I should say.

And so we do see a lot of opportunity for improvement on this measure and we have seen some strategies out there in use that can actually lead to better performance.

(Venita): Thank you. And just one follow up question, and I don't remember the details, are both of these fasting blood glucose and fasting lipid or is it - is that one of the requirements?

Woman 1: We can check on that unless (Taylor) do you know the answer to that question?

(Taylor): I don't believe that fasting glucose is required at this time or fasting lipid. I think at the development of the measure, we really wanted to give people an opportunity to start this process of monitoring. But yes, I believe that fasting is not a requirement at this time. But you know, I think we're going to go back and take a closer look at the codes.

(Venita): Okay, thank you. I was just trying to see if that was causing the lag.

(Bonnie): So this is (Bonnie) I think it's important to remember as I too was looking at it at the health plan level. So, you know, one of the other issues as far as difficulty interpreting the data from this measure is that you could have more than one, you could have two different providers, giving that (anti-psychotic) within that time period, and you won't be able to tell if that provider Is actually monitoring the medication safety or not.

Man 1: (Peter), any additional info? At the moment we've talked about a number of issues here some of which are about, about current performance gaps, so we're probably better off to try to stay on one topic at a time and so does anybody have any additional comments about performance gap specifically?

So hearing none, go ahead and use Survey Monkey to register your votes on gap and as people are doing that I will ask, (Bernadette) and (Bonnie) to move to reliability please.

(Bernadette): (OFC) was provided. So overall reliability statistics were .875 for commercial plans and .9858 for Medicaid plans and overall the majority of the committee had no reliability concerns.

(Bonnie): And this is (Bonnie), no I think our earlier discussion also flushed out the specification. And again, it's at the health point, plan level, so all they've got is the signal to noise testing.

Man 1: And anybody else have to have issues that they'd like to get on the table with respect to reliability please? Hearing none, register your vote in Survey Monkey and we'll move on to validity please.

(Bernadette): So updated validity testing was provided and the score level in both for construct validity testing indicated moderate to weak significant correlation between this measure and other measures with preventative care components.

(Bonnie): And this is (Bonnie) again. It's again, you know it's you know and secure it with convergent validity. In the corresponding measures for adolescent well care and well child visit and what they didn't know was this is the issue at the end. So did they look at convergent validity with just a numerator for glucose

because you would spin a glucose often times at the child well visit, but not cholesterol.

Man 1: That's a question for the developer I believe.

(Emily): This is (Emily) So this was done with the total rate. So looking at if they received both the glucose and cholesterol, we don't have results yet data results yet for the separate rate as those were just implemented this year.

(Bonnie): Okay. Thank you.

(Jeff Sussman): This is Jeff Sussman. I have concerns about using convergent validity with other aspects of well child care. Research that I'm familiar with and we've actually conducted shows that one area of performance doesn't necessarily track the other particularly in mental health, or monitoring of mental health medication use.

So I'm not hardly moved by the use of convergent of validity, excuse me, here.

Man 1: Thanks, Jeff. Anybody else have they have questions or concerns about the validity of this measure? Hearing none, would you please register your vote and Survey Monkey and we can move on to feasibility please.

(Bernadette): This measure has been widely implemented and again the committee really didn't have any concerns about feasibility.\

Man 1: (Bonnie), anything to add to that?

(Bonnie): Nope.

Man 1: Anybody else have anything to add? Register your vote, please. And let's talk about usability and use.

(Bernadette) It's publicly reported and it's used in current accountability programming.

(Bonnie): Nothing you need to add.

Man 1: Anything else from anyone? Register your vote, please. And any anything that anybody would like to add about usability.

(Bonnie): This is (Bonnie) I think that, you know, we only have four years of data and slight improvement, but overall no substantial change.

(Peter): The truth is - this is Peter. This is a dimension on which I have some concerns about this kind of a measure. This one seems like generally speaking, as (Bonnie) and others have that it's a relatively low bar measure.

It seems that one ought to be able to generate improvement therapy. It's not a - it doesn't seem like either a heavy lift or at least a terribly heavy lift or a measure that seems conceptually that it would take a take a long time to generate improvement on. So I do have some concerns here and sometimes elsewhere about measures that that that seemed to bump along over a long time without much improvement.

Man 1: Anybody else have comments or concerns with this one?

(Venita): This is (Venita) I knows exactly what my concern was. If we're going to continue measuring something and for three years straight, there's been nothing to demonstrate improvements just kind of beating the horse with the

same stick, I don't know if we're going to get anything different if we don't understand it.

It looks like if they tease out the two parts, and provided that data to the health plan, perhaps that will allow the combined measures and (Bertaca) is now the health plan will know what is causing the plan.

Usually health plans have all their claims data so they would have been tracking that independently. So I'm at a loss myself.

Man 1: Other- to be fair sometimes public reporting is an extra - is an additional lever. So there is that.

(Venita): Sure.

Man 1: Anybody else have usability comments on this one?

Man 2: I think the thing that our focus doesn't really touch on is the quality improvement follow up on these publicly reported measures and it's a much more complicated question to know why these rates aren't improving at all for this and other such measures.

You have been, as you suggest, Peter, I mean, this just drives the blood test is not like a real complicated task. Yet, there's probably something in that black box we just don't understand. Maybe it's people don't believe that it makes a difference and that we shouldn't treat this blah, blah, blah. Who knows?

Man 3: Or maybe it's a workflow problem.

Man 2: Yeah, they're all kinds of possible explanations.

Man 1: And anybody else have anybody else have comments they want to register on usability. Go ahead and register your vote please and would anybody like to make any additional comments that haven't already been made as a as a closing argument before you vote on overall suitability? Go ahead and register that vote too and we'll move on to 2801.

And if NTQA would like to bring this one up with any general comments that haven't been made or haven't already been made in your initial comments.

(Emily): Sure, this is Emily again and I'll be brief here, but I'll lay out that for this next measure 2801 for lifecycle social care. We're really looking at children and adolescents that have a new prescription for an anti- psychotic when they do not have a primary indication, diagnosis for receiving that drug.

And then we look to see if they have psychosocial care as first line treatment either by receiving that psychosocial care prior to the dispensing of the first anti-psychotic or very shortly after. And you know, this is really to get that confirm that we have that many children are prescribed these medications.

You know, without an indicated you use where guidelines recommend that psychosocial care should be tried and explored before in any psychotics would be considered. So, similar to the measure we just discussed, this measure is also intended for use enrolled in Medicaid and commercial health plans and again, has been reported at the health plan level for (CEDUS) since 2015 and it is also used in the Medicaid care, of course that for state reporting.

Man 1: And I'll ask the same question that I that I asked on the last one. Do you view this as a maintenance review? We - the committee has previously endorsed this measure in ’16 and do you believe that anything has either about the

measure or the something like the evidence environment around the measure, do you believe that there have been any substantive changes that the committee ought to be aware of as we reconsider the measure?

(Emily): Not from our perspective, and the only real change that has been made to the measure itself is the combining of those two lower eight strata that we noted for the previous measure.

Man 1: Thank you. So I have (Maidy) and (David) as primary discussion from this one. Would one of you just like to kick us off on evidence, please? And again, this again is a maintenance measure. So we might be able to spend a little bit less time on things like evidence and scientific acceptability and a little more time on things like use and usability. Thanks.

(Maidy): Right. This is (Maidy) (David) is that all right?

(David): Yeah, go ahead.

(Maidy): Or do you want to do it? The data were updated and that was important because for the measurements evaluation. I found it quite important that using - that the evidence shows using the - what the analyses that (NTQI) did - some significant disparities in the measure -which are quite useful at the health plan level for quality improvement purposes.

But there are no new studies that contradict the current body of evidence, I don't know that we need to go over again and again. (David)

(David): Yeah. So briefly, so this is – I'm using the parameter for use of atypical antipsychotics in children from July 2012 - so it's practice parameters from eight years ago, which means this even older, older data. So some things got

changed in terms of labels and diagnoses, what we're calling (DMDD) now, might have been called an eating disorder, I know it was back then.

And so there's some different information. A lot of my comments - I did have some critical comments that I put in there but I think it might have been with my misunderstanding.

Can we define cycles, psychosocial care as first line? What? How are we defining psychosocial care? Because I think I might have misinterpreted it as being getting plugged in with a therapist or skills worker or something along that line. But are we including psychosocial cares for psych patient as simply seeing a psychiatrist for a visit or how are we defining that?

Man 1: That's a question for yes.

(Emily): This is (Emily) and I can respond. So to identify psychosocial care, we do use a set of primarily claim codes that do include you know, your typical psychotherapy visit that you would get with a psychiatrist or other mental health practitioner.

We have family psychotherapy. posts as well, but you know, there are other types of psychosocial services such as behavioral therapy, crisis intervention, and others that are included in that and so it is fairly broad and it really indicates that there has been some type of psychosocial care provided to that child before again or shortly after they received the first antipsychotic medication.

(David): Yeah, so just to illustrate an example. So a child with fetal alcohol syndrome, severely disruptive - throwing chairs in the classroom because it's my clinic, Bill 99214 and I started to get on an atypical antipsychotic.

So will that billing code suffice as psychosocial care?

(Emily): Sorry which code?

(David): 99214. So now when I pulled up the list...

Man 4: No, 99214 is just an E&M code. You'd have to add one for psychotherapy.

(David): Yeah, but I saw 99214 in the list of codes that that was embedded in the link.

Man 4: Yes. I see what you're asking and it's a good question.

(Bonnie): So this is (Bonnie) and in that clinical scenario that we don't know when he was referred as a psychiatrist to see it, whether in the claims data there's evidence of prior psychosocial contract or any service of psychosocial care prior to that contact with the psychiatrist?

(David): Yes. It's almost as specific example as I'm illustrating, saying that there – the initial visit they get no services absolutely nothing, but as (unintelligible) can be heard at least nothing happens.

(Bonnie): Yeah. So severity drove the referral immediately to the psychiatrist. The other point is is that, you know, of course it's a data source so a limitation is that we know in claims data there's no – there's a few procedure codes for evidence-based practices for psychosocial treatment. So we get - that's a limitation. A strength of this measure...

((Crosstalk))

(Beattie): (Unintelligible)

Woman 1: ...is that it's one of the few that is pushing combined treatment for children. And, you know, saying…

(Beattie): Yes.

Woman 1: …thanks first before you. And so, you know, it has a special space since there are other targets, child psychiatric disorders that also recommend combined treatment.

(Beattie): Yes, I thought that was the critical piece for this measure.

((Crosstalk))

(David): But that's exactly my point. The measure is indicating first do psychosocial interventions before you start a med.

(Emily): This is (Emily) I'll just make a couple clarifications. So the measure does allow for psychosocial care to be provided also up to 30 days after the first prescription and that's to try to account for that scenario where you have individuals who may need more, you know, acute care or get prescribed an antipsychotic, you know, after maybe an exacerbation of symptoms. But and again the measure is focused on children and adolescents that do not have a diagnosis that would indicate antipsychotic use. So we…

(Harold): Right.

(Emily): ...exclude individuals yes that have bipolar disorder, autism, etc. from the measure.

(David): Yes, but that's exactly my point.

(Emily): Yes.

(David): So this kid doesn't have a label that qualifies, doesn't have the services yet, you work to get the kids services but whether or not that happens in 30 days is, you know, the question. So any oh…

Man: So this is…

((Crosstalk))

(David): ...(unintelligible) I think this is a, you know, an absolutely reasonable measure in things that we should do and strive for. But I guess I would advocate, you know, just I think it should be worded a little bit differently than if, you know, for any label, any situation where were a kid requires, you know, that type of medication that we should be pushing for those psychosocial interventions regardless of label.

(Peter): (David) didn't you say that - sorry I – that just of the code numbers don't actually translate to me in my workstream? So if it - didn't you say that it's a code, your billing code would count as the psychosocial care or that it wouldn't count as a (unintelligible)

((Crosstalk))

(David): Yes, so 99214 looks like it's listed. But then so if there's even primary care and they do a 99214 is that going to count for them too?

Woman 2: Can you clarify where you're looking at the list? I'm having trouble figuring out where the value set is? But I see one document but it's a whole bunch of codes.

(Peter): And (David) it could – if the issue is that there might be some things that pass this measure that we don't think sort of represents depth and care that's sort of like then we sort of discussed that on the last measure too it's that even at - even given that the performance on this measure has a fair amount of room to move. And so…

((Crosstalk))

(Beattie): Right.

(Peter): ...right? So we'll need - the question that's on the table right now is do we believe that there's evidence to try to essentially to try to encourage the provision of psychosocial care as well as antipsychotic meds?

Woman 2: Well I think before that there's just technical question about what is defined as psychosocial care. Is it E and M code alone psychosocial care? Could the developer clarify that?

(Emily): Yes, this is (Emily) So that 99214 code is not included in the psychosocial care values set that we used to identify psychosocial care. You did probably see it though because it's one of the codes that can be used to identify when someone has a visit for one of the exclusion diagnoses. So it is not – it would not count alone as…

Woman 2: It would not okay.

(Emily): ...psychosocial care, yes.

(Beattie): That's a point. Thank you for clarifying that.

(Peter): So would anybody else like to on the evidence criteria would anybody else like to raise any issues that haven't already been discussed?

Woman: (Unintelligible)

(Andrew Sperling): This is (Andrew Sperling), I apologize but I did not – when I was looking I did not see the exclusionary diagnoses when I was reviewing it the first time. Can someone at NCQ clarify that please just briefly or was it - could it be any psychiatric diagnoses or was it like specific…

Woman 2: No.

(Andrew Sperling): ...to, you know, first steps of psychosis or mania or...

(Emily): So it is - this…

(David): (Emily) can you…

((Crosstalk))

(Emily): Yes. It is specific to diagnoses where antipsychotics may be indicated for children or adolescents. So just quickly the types of disorders that we exclude for example would be schizophrenia, schizoaffective disorder, bipolar, psychotic – other psychotic disorders and autism and I believe tic disorders as well.

(David): So I didn't see tic disorders in the…

(Emily): Yes I believe it's in the developmental disorder value set.

(Beattie): Yes.

(Peter): Anybody else have questions or comments or concerns on the evidence criteria and please?

(Bonnie): This is (Bonnie) I thought that was a strength of this measure that NCQA did a pretty good job excluding SMI.

(Beattie): Yes they did.

(Peter): Thank you. Any other comments? All right I'm going to try to move us along. We're down to about 15 minutes and I'd really love to finish this measure. So please…

Woman: Yes.

(Peter): ...vote on the evidence criterion please. And could we move on to performance gap please?

(Beattie): This is maybe there's room to improve, lots of room to improve.

(David): Nothing to add.

(Peter): And so for people who want at least ballpark numbers we're talking about 55% to 60% performance.

(Beattie): Right 55 to 60.

(Peter): So anybody else want to comment on performance gap, register your vote please. And let's talk reliability please.

(Beattie): Okay well I haven't got much to say about that other than that the distribution at the health plan level reliability showed that half of the commercial, all of the Medicaid health plans succeed at .7 thresholds for liability.

(Peter): Thanks and (David) anything to add?

(David): Yes not really. I think you referred to it as moderate and I don't disagree.

(Peter): Anybody else have comments on reliability? Register your vote please. And let's move to validity please.

(Beattie): Construct validity was tested using four measures that also require a high level of coordination across settings and providers which is the teeth of this measure. Follow-up at the hospital or ER in seven or 30 days for children and adolescents. Now leases indicated that the measures have parlayed in plans that have higher rates on one will have higher rates on the other suggesting the measure had adequate validity. That's it.

(Peter): Thanks (Beattie) (David)?

(David): Face validity and empirical validity testing NQF agrees with it. They only - so in terms test of validity I don't know if this is where it would fall but just in terms of, I do think that they may be under other clinical scenarios where it really would be appropriate to move forward with medication in conjunction with trying to get psychosocial services and therapy and other supports going.

But to say that you should hold off on potential effective treatment that could relieve a lot of suffering and help turn things around more quickly in other if we want labels thing like DMDD or kids with bad brains, fetal alcohol, kids who were meningitis when they were young or and actually brain events or other structural or bad brain reasons where there may be certain situations where, you know, it really could get things moving better more quickly. I just have some concerns in specific situations where you don't want to hold off on something that could help.

(Peter): And this is you - this is one of the many reasons on essentially every measure where there are sometimes good reasons that the right target number isn't 100%. Anybody else have comments on validity? Let's register your votes please. And feasibility please.

(Beattie): Well the state elements are in defined fields and electronic claims. So and this is not an e-measure. It's publicly reported in nine different applications, seems eminently feasible.

(Peter): And (David)…

(Beattie): Yes. I'm kind of…

((Crosstalk))

(Peter): ...do you have anything to add to evidently feasible?

(David): Electronic claims currently in use medicated course that nothing else to add.

(Peter): And queue up I think we need a new category going forward.

(Beattie): This is (unintelligible)

(Peter): Anybody else have – anybody else have any comments or questions about feasibility that haven't already been addressed? Hearing none, please register your votes. And (Andy) or (David) if you'll move us to usability and to use please. So use first.

(Beattie): Well for - yes primes indicate no significant barriers to implementing the measure despite the need for coordination of care which often is cited as a barrier. They have not - they say there's no - not any particular barrier so it should be quite – and it's being used so the gap is pretty bad.

(David): Although my opinion isn't perfect I, you know, I do, you know, think it there is a pass. But again I do think that there might be better things to measure that would remove the labels and just say if a person is on these types of medications regardless of label that it's worth making sure there's some other people working with the kids and families good care coordination and so forth.

(Peter): But we always have to vote on the measure that's in front of us but NCQA please, please take that under advisement for thinking about things going forward. Anybody else have comments on that issue?

(Beattie): My only comments on that issue...

((Crosstalk))

(Beattie): ...one little comment is that this measure importantly focuses on something that is quite serious. I mean antipsychotics are serious medications. I know any medication for a child or adolescent is but I am glad this is the focus.

(Peter): Thank you.

(Andy Ripp): This is (Andy Ripp) (unintelligible) (Alex) said one thing at least in terms of what I found interesting about this is that, you know, prescribing antipsychotics that's the first step of psychosis was validated in the NIH (RAISE) study of coordinated specialty care but it was the low dose antipsychotics with very, very careful monitoring. And these are, you know...

((Crosstalk))

(Andy Ripp): we're largely talking adolescents here that have already experienced a psychotic episode. So I agree with (Beatty) on this. It's a serious thing and this type of thing is, you know, when it's done it needs to be very, very carefully monitored.

Man: Yes.

(Peter): Thank you. Anybody else on use? Hearing done can we move to usability please?

(Beattie): As I've said the plans did not find that there were significant barriers to using the measure.

(David): Nothing new to add.

(Peter): My only comment to add as a feeder is that this is another one of these measures that doesn't seem to be - and I think it's important too and it doesn't seem to be showing a lot of improvement so probably…

(Beattie): Yes.

(Peter): ...the world – I'd love to see the world A, come up with ways to generate more and faster improvement on this important issue.

Man: Yes.

(Peter): Anybody else have…

((Crosstalk))

(Peter): Does anybody else have closing issue – comments on this issue? Hearing none please register your votes. And then I'll just ask it generally if anybody has a closing argument on this measure before we vote on overall suitability?

(Andrew Sperling): This is (Andrew Sperling) One last – wouldn't call it (blight), but a concern and that is I think it's been four or five years since the Iowian study came out on fidelity to the models we know for psychosocial services in psychiatry and found, you know, sort of abysmal results in terms of people out there, you know, that they're doing CBTs, they're doing these different things but fidelity of the evidence-based model in terms of how it's actually done in practice the findings they were quite troubling in terms of the fidelity. And, you know, it's one thing to check a box and say you're doing a psychosocial intervention. It's another completely different thing – and I know this measure probably isn't even intended to follow this but whether or not the practitioner is – if there's fidelity of the model.

(Beattie): Yes.

(Peter): Yes.

(Beattie): You're right.

(Peter): And anybody else have a closing argument that they'd like to register? Hearing none please register your votes. And I – and I think at this point we've completed our measure. Thanks for sticking with us while we power through those and I'll hand it back to (Nicolette) at this point.

(Nicolette): Thank you very much (Peter) and thank you to everyone for working with us as we switched voting platforms and also making sure that we got through all the measures we needed to today. We really appreciate it. So that does conclude our measure discussions. We do encourage you hopefully you have been using the Survey Monkey link to capture votes as we're going along. Please do remember to submit the survey at the end so that we record your responses.

We will also be sharing the link with the other committee members that either had to leave for a portion of the call or were not able to join us today along with the recording in order for them to cast their votes. Another comment just before we conclude the call, we did originally plan a discussion around harmonization of related measures. I did want to note NQF staff did look at measures to see if there were any related or competing measures. We did note that there are no directly competing measures for any of the measures that are up for either new endorsement or maintenance endorsement.

And the measures that are related we did feel that there was justification for the differences, either differences in target population focus areas, level of analysis as a few examples. So we didn't have any major concerns to point out but we did want to see if there were any concerns from any of our committee members. We don't have time to have a full discussion today but if there are

any thoughts on harmonization that you would like to discuss, we should have some time on the post comment call to pick up that discussion. So – but I did want to note there were no directly competing measures.

(Peter): And maybe as a process suggestion since we're at our last five minutes if folks have issues that they'd particularly like to address on that call why don't you send staff an email and we can – and the chairs of the staff can work on inserting those issues into the post comment call.

(Nicolette): It's a great idea, thank you. Thank you (Peter) We would appreciate that. I think with that we will go ahead and open for public comment. So if there are any members of the public that would like to make a comment at this time your line should be open. We encourage you to do so. You can also send us a chat through the Webinar platform if you would like to make a comment at this time. Okay, hearing no members that want to make a comment at this time we'll go ahead and move forward. I'm going to turn it over to (Hannah) to share a few next steps.

(Hannah): Thanks (Nicolette) Thank you again everyone for joining our call today and for staying engaged. We have a couple of next steps. We will be sending out the voting link to the other committee members who either had to drop off early or who could not make the call. We are giving them 48 hours to listen to the recording and cast their votes. Once we reach quorum on all of the measures through the survey link we will send an update with voting results to the committee members and to the developers just so you know what the outcomes were. And then staff will work to draft a report that will summarize our discussion and the outcomes of each measure. And then we will be posting that report for public comment from March 11th through April 9th.

Once the commenting period closes staff will update the report and we will reconvene with the standing committee to discuss any comments that were received on any of the measures and to revote on any measures that where consensus not reached. And so it is important that we get a quorum on that call. And it's on April 22 from 12 o'clock to 2:00 pm Eastern Time. Calendar invites have been sent out for that so if you have not received it please let us know as soon as possible. Again it is important that we reach quorum on that call so that we can revote on the measures. Otherwise we'll have to do the Survey Monkey again.

And then after the post comment call we will finalize the report and it'll go through our consensus approval committee, our (CSAC) committee who will make the final endorsement recommendation for each of our measures. And then after that any measures that are passed will go through an appeals period for 30 days and at that time any member of – any NQF member or member of the public can appeal any measures that were (unintelligible) And so that's it for the fall 2019 cycle.

A couple of dates for the spring 2020 cycle. Our intent to submit was on January 7. And so far we are expecting three measures for the spring 2020 cycle which I would see on the slide here. Staff will once the measure submission deadline passes staff will go through the same preliminary analysis process and will keep you updated on any next steps for the spring 2020. That's all I have. Does anyone have any questions? Please for free to let us know either now through the chat or via email. Our project info or project contact info is listed on the slide here.

Hearing no questions I just wanted to say another thank you behalf of the NQF staff especially (Harold) and (Peter) for your leadership on these calls but all of our committee members and developers for the work that goes into

this and your time to be present during the discussions and contributing and answering questions. We very much appreciate it. (Harold) or I think (Harold) actually had to drop off. (Peter), do you have any final…

(Harold): Yes.

(Hannah): ...or concluding remarks?

(Harold): Actually I'm still here and…

(Hannah): Okay.

(Harold): ...and thanks again to NQF staff for getting us through this and also to my co-chair (Peter)

(Peter): And back at you (Harold) Thanks to all the committee members. We went through a lot of measures this cycle and including some hard ones. And thanks for hanging with us. I also want to particularly thank the staff. They always do yeoman's work and this time you won't know this unless you're an NQF aficionado. But I know this that there's been a lot of recent comings and goings at NQF. And so their degree of difficulty has been harder than it sometimes is. And so I - we should particularly express our appreciation for them doing their usual stellar work under what were probably harder than usual circumstances. So thanks for that.

(Hannah): Okay thank you all. We will be following-up with everyone and we hope everyone has a great afternoon. Thank you.

Woman: Thank you.

(Harold): All right thanks.

((Crosstalk))

(Harold): Thank you.

END