NATIONAL QUALITY FORUM Moderator: Benita Kornegay Henry 06-21-19/3:25 pm CT Confirmation #21926731 Page 1

## NATIONAL QUALITY FORUM

## Moderator: Benita Kornegay Henry June 21, 2019 3:25 pm CT

Man:	But I'm on the web now.

Erin O'Rourke: Okay. So if everyone, it's sounds at least we've got everyone on the phone and you can figure it out the web. So why don't we get started in the interest of time, we've got quite a bit to cover today.

> Good afternoon everyone. This is Erin O'Rourke, Senior Director with the National Quality Forum. Thank you for joining our Second All-Cause Admissions and Readmissions Standing, Committee Spring Cycle 2019 Web Meeting.

As a quick recap of yesterday, we started our review of Measure 2539. We still have some outstanding issues to discuss before we jumped into the official agenda for today.

So I think with that, why don't I turn it back to Christie Travis our Co-Chair to reopen the conversation on validity. I think in particular, the committee had a few questions they wanted to continue discussing around the risk adjustment model of Measure 2539, the outpatient colonoscopy hospital visit measure and the potential needs for adjustment or social risk factors.

Christie Travis: Sure. Thank you, Erin. I appreciate it and thank you all for coming back for day two and really appreciate what we were able to accomplish yesterday. We did want to be sure that we left enough time for a thorough discussion of the validity aspects that Erin just mentioned for 2539. And I think where we were was to really begin our in-depth discussion of the social risk factor adjustments.

Just to kind of refresh our memory and to get us back on a common playing field I was going to ask Tony and Keith if you all would be willing to kind of share with us your comments as we discussed this around the social risk factor adjustment. And I don't know for sure who all is on the phone. I know Keith is. Tony, are you on the phone?

- Tony Meoni: Yes, I am. Hi.
- Christie Travis: Would you mind kicking this off?

Tony Meoni: Sure. Sure. So the question about the risk adjustment for the social demographic factors had to do with the fact that there was a difference between facilities or patients who had dual eligibility, which is their measure of the sociodemographic factor and non-demographic eligibility status depending on the facility type. And the difference was about 2% to 3%.

So I think that we just left it. I had asked question whether or not when they did the three-year analysis, if the risk adjustment of the - using the dual available status changed the C-statistic at all. I wasn't sure in the write up

whether or not they were reporting the initial one-year result or the three-year result for that.

And then we sort of stopped there in terms of making, going into more detail about including it - not including it in the model. So I think that was where we ended up.

- Christie Travis: Thank you I think so as well Keith, any opening comments you'd like to make about this issue?
- Keith Lind No, only that I was referring to one of the reviewers on the scientific panel, who said that the dual effect was significant and consistent across facilities, which, to me made it seem important. But if and the reviewer pointed out that if you're not going to adjust for dual status then you either adjust for something else. And I understand that we're only supposed to vote based on the data that's been submitted. And that's where it stands I think.
- Christie Travis: Yes, thank you for reminding us of that issue as well. And Erin I don't know who's on the phone, is Elizabeth on the phone and maybe can you talk about this a bit of your analysis and whether or not the dual eligibility results changed since the meta-analysis was completed.

Elizabeth Mitchell: Hi, Christie, it's Elizabeth.

Christie Travis: Hi.

Elizabeth Mitchell: How are guys. Thanks for continuing to consider this competitive measure. We had, I talked to CMS yesterday because we're just going to get back to the dual question that this issue of, you know, what can we share from the three-year data.

And I was offering yesterday, for example, sharing the distribution of the measure scores and outliers but we got off the call. I talked with (Vanina Meyer), who's the contract officer and presenter, the supervisor of this work at CMS. And Dr. Meyer was, she's very, she's very pragmatic sort of person.

So we were just thinking, in general, just getting back to the dual question, which I promise I'll answer. It's hard for us and for you guys, if we proceed sort of piecemeal giving you the three-year results. We mentioned yesterday, we don't, it's not fundamentally different. But one reason that CMS went to this three-year result is that we can see everything better. We can see differences and performance better. We can see disparity better because we have more cases per facility.

So we have, we can make more equality inferences and we have better reliability as you saw on the numbers we shared. So CMS went to the threeyears for that reason but as a practical matter I wanted to check in with the committee and NQF staff.

I'm not sure, it just was a process measure, it's helpful for us to sort of trickle out those results. And ask them if we're doing work under the Impact Act right now. We haven't quite wrapped that up that would be done in July, sort of late July, mid-July.

Looking further at disparities issue, so I'm struggling with how to proceed and what (Vanina Meyer) suggested is probably we should try to stay in the lane that we're in, which is we have a one-year measure in front of you with application with all that data. We did provide supplemental reliability numbers those were published, CMS published those during rulemaking. So probably it's hard, it's not that easy for me to trickle this data out. And I don't think it'd be easy for you to give you sort of part of, but not a full picture of the three-year data.

So given that I was going to ask NQF staff, kind of what they recommend here. I guess I'm just saying, we because of the transition of CMS policy and how to use this measure of credible measure was already in front of NQF. I'm just wondering how we should proceed.

I think CMS would say listen for with a one-year application, if that's not, that's not going to work. What are the other options to get the committee what it needs rather than sort of halfway going down that path.

Erin O'Rourke: Sure. This is Erin, Christie, if you're okay, if I jumped in here?

Christie Travis: Of course.

Erin O'Rourke: Thanks for that that's helpful. So I think at least from the NQF perspective, I think we have a couple options on the table and maybe we can see what the committee feels comfortable with.

I think the application is complete for the maintenance of the measure with the one year time period. So I think the committee could continue to consider as submitted and pass except that the three-year data is supplementary in an attempt to try to synchronize what the committee has available with how the measure would be applied in the program.

I think recognizing that change has been made and wanting to be upfront with everyone about the application of the measure. And that CMS will be using the three-year time period for the ASC QR and the OQR programs where the measure is being used.

I think we also have an option. And again, I don't know if we want to go there right now. So the developers to withdraw the measure from this cycle and bring it back to the committee as - with the updated data for a three-year for the measure to be reviewed with the three-years' time period.

So we may be stopped for now, if people are feeling, this is a little too confusing. And we kind of put it on hold and come back to this measure in the fall or potentially in the spring depending on when data would be ready and review it with the three-year period and maybe not get into the one versus three-year question.

So I think those are the options and maybe if people feel comfortable looking at the measure as written with the one-year time period or people prefer to maybe put a pin in the conversation for now and come back. When CMS (unintelligible) are able to update the forums and we can look at it as a threeyear measure.

Christie Travis: This is Christie. I'm just asking kind of a clarifying question. What are the implications or what are the possible implications of taking the second recommendation? And in I mean, are there any kind of timing issues that would be missed?

I mean, CMS could put their measures into their programs, whether they're endorsed or not. And I guess I'm just trying to understand what the disadvantages of going with Number 2 would be, if people would like to articulate that. Erin O'Rourke: Sure, I can start with from the interest perspective and if Elizabeth has anything to add from the developer perspective. From the NQF perspective, measure developers have the option to withdraw their measure from this cycle basically anytime up until it goes to feedback and work with us to resubmit to a future cycle.

So there'd be no, you know, penalties to the developer from the aspect of, you know, as a measure would not immediately lose endorsement. It would stay as an endorsed measure and we'd work with them to reschedule the maintenance review.

I think potentially, there could be some challenges for the developer with timing, contracting, budgeting. Obviously, it takes some effort to update the forms and come back to NQF. So I don't want to minimize anything that, you know, may impact them on that aspect.

So from the NQF perspective, the measure would stay endorsed until we either – no, it times out on how on the grace period for maintenance and placement or we have that final feedback decision to remove maintenance endorsements.

So there'd be no penalty for the developer and the measure wouldn't immediately lose its status and we'd work with them to reschedule and bring it back to you.

Christie Travis: Great that that's very helpful. And if I understand correctly, it's really the developers -- the developer would be the one to withdraw it as I understand what you're saying.

- Erin O'Rourke: Correct. We either NQF for this, the committee doesn't really have the ability to like, to get out and process at this time. Right. Okay.
- Christie Travis: So I don't know if there's anything to share with that Elizabeth or was that somebody trying to come in.
- Leslie Kelly Hall: Sure. I think it's Leslie, I just had some questions about and it's about usability. But the three versus the one year on a maintenance measure, with one that has a clear use case and outpatient. It doesn't seem to be a lot of ambiguity around how to select this patient type into this procedure.

So I'm considering what the provider implications are for one year versus three years on a maintenance measure, is it more difficult, less difficult to affect change and have it reflected in the outcome if it's one or three years?

Christie Travis: Yes. Well, I don't want to, I would like to kind of - lot of keep us thinking about these two options for the moment.

Leslie Kelly Hall: Sorry.

Christie Travis: No, that's okay.

Leslie Kelly Hall: Okay.

- Christie Travis: I mean that's a good question but Elizabeth, I didn't know if you had some thought really options.
- Elizabeth Mitchell: I can just say quickly, Leslie to your question that was one of the considerations in CMS, you know, vetted in the rulemaking process about

going to three years. Because we all, you know, we quality measures, you know, people using measures being measured.

We all want to have a shorter time frame and that's the trade-off when you have a low volume outcome to and, you know, variable size facility. Some of these facilities do many, many medical uses, some do fewer.

So yes, if that - that was waiting the decision go to three-years the decision go to three-years has been made by CMS and finalized. So back to the point that you're asking Christie, I think, the spirit of my conversation, you know, I know it's the "Developer decision," but as Erin mentioned, the cause really falls on CMS. I mean they would, you know, their staff and with support from contractors like they would put up data form.

So, but the spirit of any of those in my discussion yesterday, you know, is just CMS would like to do what makes sense here and I think it's a committee and we've all investor work in this. We, you know, have the developer CMS has and you all had in reviewing it. If you want to move forward with a one year and, you know, finished endorsement maintenance process and you think that we can do that without a lot of confusion. We're fine with that.

But if it fits at my own experience of the conversation is you all are asking a lot of good questions. You know, you want to see the data, CMS will have the data because as I mentioned yesterday, they're putting the measure into reporting the three years of data, it'll be public in that reporting result will be published in early 2020. So they're pulling the data now to generate those measures for results.

So I have the capacity that it doesn't post a constant number they had the capacity then to (unintelligible) form with a three-year results. So if you

thought it was important, I think CMS essentially said, if it makes sense from the point of view the committee, certainly CMS and definitely, you know, we as a current contract around this are fine. Just waiting, updating the data and resubmitting it.

And so that's why I wanted to raise it with all of you, although it would be our decision on meeting CMS is really and, you know, together, I think just wanted to vet it because if it's, it's hard for me, you know. As I was saying, like, if we're giving you these pieces of analysis, for example, if I gave you the outlier numbers and the variation numbers.

We sort of, we have those from a bit of a patch and gather dataset it's, you know, it's not a complete three-year datasets with the same breadth of information we couldn't use it to do that, some of the analysis.

So to really give you all the three-year results one common dataset, you would have to wait till the dataset is available that's getting generated right now for public reporting in the winter. So that's why it seems I know as somebody who, you know, when I'm doing research, I don't and I think they are coming from lots of different places with a lot of caveats becomes kind of...

Christie Travis: Right.

Elizabeth Mitchell: ... to challenge. So that's - so I'm trying, if the committee says, whenever we're good enough with this one year of data and we're not, you know, and we have the reliability numbers which I actually think is the biggest issue with using the year is the reliability.

But beyond that you're gaining to go forward. We are too but if it's too much, and you really need a three year, those three-year results, then I think we

should consume that they would be comfortable referring to, you know, the next cycle or...

Christie Travis: Okay.

Elizabeth Mitchell: It might, because that's in August, it might be the next cycle after that because it's a very, very close deadline.

Man: Sure what this is Travis one question for staff. There's a lot business metrics that are based on rolling three-year averages. So my question would be whether there might be a way to ramp up from one to three years with this measure, with the understanding that we're creating a three-year baseline through the first three years of individual reporting.

> And you know that's that may be apples to oranges but there's tons of business metrics that are based on three-year rolling averages where you drop off, you know, the trailing year as you accrue more data as time goes by.

Elizabeth Mitchell: And that's what it should be. I mean, it's a wrong three-year average. But the steps to using it is going to, it's already set. And so that's not a decision before the committee's decision and so cognition have the value of your time and your, you know, your thoughtfulness.

> So I just, it just do you think we should wait. It would have been easier for the committee to consider this measure if we bring it back to with a three-year dataset and all those and all the analysis running for three years of data.

Erin O'Rourke: And this is Erin. I think I'll just to jump in that we and I think it's a question for Elizabeth of whether that's the conversation is helpful in preparing that application. You do have - that the process allows the withdrawal up until the CSAC meeting. So there's we, it doesn't need to be made today if the conversation is useful to you to prepare for the future or if the committee feels comfortable with the one year data.

As a reminder, this was endorsed initially with the one-year time period and that was how it was implemented. So if you're comfortable with the measure, you know, maintaining it as it was or if it's Elizabeth finds this useful for preparing the application for the future we could finish the conversation and go from there. But if everyone's too confused, we could move forward with the withdrawal.

Christie Travis: All right. This is Christie, I guess. And this is I'm taking off my chair hat now and just commenting as a committee member. What I'm struggling with is that the way that the measure is going to be used as a three-year measure. And, you know, to focus on a one-year measure that isn't going to be used at least by CMS. I guess it could be used by other people.

But, you know, it's not going to be used by CMS seems not to really be very satisfying and I don't mean personally satisfying to me. But it almost seems like it's jumping through a hoop that really is not going to necessarily advance, you know, advance the work. And that's why I'm struggling with it. And the other piece that if I'm and I think I'm hearing this pretty loud and clear is that this is going to be is a three-year measure.

And so, you know, just in my own mind, I struggle with going through an endorsement from for a measure that's not going to be used. So I just lay that out. I'd be glad to hear obviously from other committee members.

- Tony Meoni: This is Tony, I agree with you. I just think it makes more sense if there's no negative in postponing. I think it just make sense that we have all of the data related to the three-year period.
- Leslie Kelly Hall: This is Leslie. I struggle with every measure we have asked for review on social determinants and/or risk factors or call, you know, call it anything there. We come back with no material finding. It doesn't necessarily impact.

So I don't know if that's the way we're collecting the data or what. But I, so I struggle with, do we, are we putting extra hoops in here without any expected change? One question and then I also agree with Christie approving a measure that's not going to be used in you know not very useful.

- Christie Travis: Other thoughts? Thank you both.
- Keith Lind: This is Keith. I agree it just seems like we should be looking at three years of data and the fact that I couldn't find the statistics that were based on three years of data sort of drives that point on for me. And you know it's been really helpful to have that for all of the measures, all of the elements here. And particularly where the tools are concerned that I mean, they noted that the sample sizes in one year phase of the number of tools were so small.

And maybe with three years of date we'd get a better, stronger picture of what the impact of the tools was. I don't know whether it would make a difference. But one of the reviewers asked what about, you know, they're only 4.2% tools in the ASCs. And what happens in facilities which is more than 50% tools, I mean wouldn't that make a difference and maybe three years of data would give us a better handle on something like that.

Thomas Smith: Thank you, Keith. I agree.

NATIONAL QUALITY FORUM Moderator: Benita Kornegay Henry 06-21-19/3:25 pm CT Confirmation #21926731 Page 14

Christie Travis: Other others on the line?

Thomas Smith: Yes. Tom Smith, I agree, given the questions that have been outlined as long as there's no significant burden to the developer or to the CMS program. I would vote for waiting, getting the three-year data.

Carol Raphael: And this is Carol Raphael and I kind of incur with the direction that we're headed is to wait unless there's any significant downside.

Thomas Smith: Thank you.

- Erin O'Rourke: So this Erin from NQF. So Christie, maybe a pass forward is, it sounds like NQF and the committee are making a strong suggestion to the developer that it would be less confusing and highly preferable to look at the measure as it will be implemented in the program and perhaps withdraw with the one-year application and come back in the fall of the spring with an updated application with the three years of data.
- Christie Travis: And I think you're reading us correctly, Erin, unless there's somebody who hasn't shared and want to be sure they have the opportunity. So if we move in that direction is the recommendation kind of just to go on and move on to the next measure. Erin from the team, get back to those...
- Erin O'Rourke: I think it was just checking with Elizabeth, just to make sure they're comfortable making that decision today. And they don't want us to the push the process just so that, you know, it's easier to have a conversation and...

Elizabeth Mitchell: Yes, I mean, I think we are again, you know, CMS shares the mental gymnastics that you were to go through with the current approach. So and,

you know, it creates uncertainty and for you, so I think we can presumptively say that well CMS hold it and resubmit it in a subsequent cycle specified as a three-year measure.

It has been really helpful to hear the questions if there are any other comments or questions that, you know, people wanted to get on the table before we move up the measure. Then that will help us explain for what we bring back. I can't promise we can do every single analysis people want.

But it's helpful to have had this discussion. So I think we would be able, when we came back to this community with a measure to pick up where we're, you know, wherever leading off as opposed to start all over.

Christie Travis: Well, thank you for that Elizabeth. You know, one that I know that Keith just mentioned that was briefly talked about yesterday was understanding the impact of the duels. Not just on the total but looking at, you know, those facilities that have a significant number of duels.

And I think you did mention that in your opening remarks yesterday but, you know, I think, you know, I think that would at least be something I would want to reinforce the Keith said from taking mature head on and just put in my personal comments.

So, you know, I don't know if there are other short kind of comments that people would want to make that could be helpful as they think about the application for the three year period.

Karen Joynt Maddox: This is Karen. I'd just jump in and say that some of that information is provided. So on, it's on Page 70 of the...

Christie Travis: Oh, thank you.

Karen Joynt Maddox: Right. No, that's not, sorry, that's the wrong measure. It's on Page, sorry, it's on 70 for the moment we haven't talked about yet.

Christie Travis: It's okay.

Karen Joynt Maddox: Let me see if I can find it in the other one. Otherwise, I'll just bring it up in the next time that they say the range of movement among groups, when adjusting. I don't know if there's just a way to maybe highlight that in the resubmission, if that's of interest to people. But they have, they do have the information...

Christie Travis: Right.

Karen Joynt Maddox: ... (unintelligible) I'm sorry I was looking at the wrong form.

Christie Travis: Thank you. Well any other comments that we want to say now so that it could be helpful in the resubmission? Okay, well, I'm hearing none. Elizabeth, thank you so much for yesterday and today. And I want to thank the NQF team for helping us, you know, kind of get down to what the options are. And I will say we look forward to seeing the measure resubmitted as a three-year measure.

Elizabeth Mitchell: Thank you all so much for your thoughtful review.

Christie Travis: Thank you. And Erin, I'm going to, I think turn it back over to you.

Erin O'Rourke: Okay, perfect. I did want to jump in. Before we start with 3495 and take us back to just do a quick roll call to make sure we know who's on the phone.

So if you're scrolling back to yesterday's slide said bring up the – so yes. Apologies, yes that's the list so I can do a quick roll call to know who's on our line. John Bulger?

- John Bulger: Here.
- Erin O'Rourke: We heard Christie. Katherine Auger? Jo Ann Brooks?
- Jo Ann Brooks: I'm here.
- Erin O'Rourke: Helen Chen? Susan Craft?
- Susan Craft: Here.
- Erin O'Rourke: Wesley Fields.
- Wesley Fields: Here.
- Erin O'Rourke: Steven Fishbane? Paula, let us know she would not be joining as did Larry. Anthony Grigonis?

Anthony Grigonis: Here.

- Erin O'Rourke: Bruce Hall? Leslie Kelly Hall?
- Leslie Kelly Hall: Yes.
- Erin O'Rourke: Paul Heidenreich also let us know about a conflict. Karen Joynt Maddox?
- Karen Joynt Maddox: Here.

NATIONAL QUALITY FORUM Moderator: Benita Kornegay Henry 06-21-19/3:25 pm CT Confirmation #21926731 Page 18

- Erin O'Rourke: Keith Lind?
- Keith Lind: I'm here.
- Erin O'Rourke: Paulette Niewczyk? Carol Raphael?
- Carol Raphael: Here.
- Erin O'Rourke: Mat Reidhead had let us know about a conflict. Pam Roberts? Derek Robinson also let us know and then finally, Tom Smith?
- Thomas Smith: Here.

Erin O'Rourke: Okay. Thank you, everyone, apologies for taking us back there so I will just move us back to our next measure under review. Perhaps in the interest of transparency, we should go to a quick public comments on anything that has been said about 2539 in the committee's process.

So why don't I just quickly pause since we had a public comment scheduled. And see if there's any comments people would like to make about Measure 2539 and the committee's decision to work with the developer to resubmit in a future cycle with three years of data.

Okay. So hearing none, in the interest of time, why don't we move into our second candidate measure? This is Measure Number 3495. This is the Hospital-Wide 30-day All-Cause Unplanned Readmission Rate for the merit-based incentive payment system eligible clinicians and clinician groups. Again this measure was developed by your core and CMS is a steward. It's obviously an outcome measure using claims and enrollment data.

The level of analysis for this measure is either the clinician group or practice or the clinician individual. This measure was reviewed by the scientific methods panel. It did pass the method panel review.

So I think with that I can turn it over to John and Christie for discussion. And our lead discussions here are Karen Joynt Maddox, Wesley Fields and Leslie Kelly Hall.

Christie Travis: John, do you want – John, do you want to leave this or what's your preference?

- John Bulger: Sure. Was that Karen ready to start?
- Erin O'Rourke: Oh no, that was Erin. I was I just going to say (Fred Lee) just as we did yesterday, focused on important first, even though we don't have quorum and we will be voting it might...

John Bulger: Yes. Okay. One of the primary discussions I like to play this out.

Karen Joynt Maddox: All right. This is Karen. Although that wasn't me before I will do it now. So this measure is looking at unplanned readmissions and the merit-based incentive payment system for eligible clinicians and clinician group.

And in the reclassified version of the hospital level measure that we're all familiar with in terms of the hospital wide archive, unplanned readmission so it's that measure we've talked about before that has the five cohorts and then combined those cohorts into a final for a full, I guess, group to evaluate 30-day readmission rates after a hospitalization.

You know, I think the rationale for focusing on readmissions has been well explored and well explained. There are certainly things that can be done to reduce readmissions, perhaps not as simple as interventions for some of the other measures that we look at.

But there are data to suggest that there are ways to reduce readmissions. And actually, a lot of the good data suggests that better outpatient management is the best way to do it as opposed to anything that happens in the hospital and in terms of intervention.

So this may actually be a little more tightly linked to services that could be provided by their providers sort of on the hook. I don't know if that goes in importance but they do show a significant range of readmission rates from 5% to 38% for clinicians and 7% to 25% for clinicians group suggesting goes a high rate of readmission and a significant amount of variability.

John Bulger: Any other discussions have a comment?

Leslie Kelly Hall: Yes, this is Leslie. And the review has some common themes throughout, it started in the initial importance to measure under evidence. And there were discussions about the week evidence in a relationship to individual submission to the health system.

And this seemed to pepper the whole measure the assumptions around the relationship with the individual physician to the care. Especially in systems where there's a high degree of hospital list and so there was I think it's worth understanding a little bit more about that.

There was also some discussion about is this really a quality measure or is this simply a new measurement of a financial model effectiveness. And I think there was some, there's concerns there.

I share that concern because the assumptions made on any nips or mock right is the financial incentives are aligned and that actually determines helps to determine a proper care. But this measure could actually, this is somewhat disassociate the physicians who are caring for the patients and the physicians who are being measured.

So there wasn't. I agree that the evidence shows that there is a weak relationship between the individual physician so I would like to hear or have discussion on that.

And then the performance gap also there was discussion, they didn't see that it was present in a submission that there was a difference of performance between one financial model and another fee-for-service versus limits the macro model or ACL model. So I'd love to hear comments on those.

Erin O'Rourke: And John this is Erin. I am deeply sorry to jump in here but I should have asked Yale team to go first to present the measure and perhaps they could while they're making their overview comments address some of Leslie's questions.

((Crosstalk))

That was my fault. I jumped up to the discussion so I deeply apologize to the team from Yale and...

(Crosstalk))

NATIONAL QUALITY FORUM Moderator: Benita Kornegay Henry 06-21-19/3:25 pm CT Confirmation #21926731 Page 22

## John Bulger: Yes.

(Jeff Haran): All right. This is (Jeff Haran), form the Yale team. And I thank you for the opportunity to present this measure. And I will try and give a quick overview of the measure as opposed to just couple of the points that Leslie raised.

Just to give a little background this is, you know, as Karen said, this is an adaptation of existing hospital wide all-cause unplanned readmission measure, which is used in public reporting for at the hospital level. It has there is we feel like evidence that clinicians who care for patients have some influence over the readmission outcome. And I'll get circle back to some of the evidence a little bit.

And as a consequence, we have adopted that measure to measure clinician and clinician groups. One novelty of our adoption is that rather than attributing patient discharges to a single clinician we have identified up to three different clinicians for each discharged that may be responsible for the patient outcome.

And as I said we expect this measure be used in the Med Quality Payment Program. So there is some thinking to that that the incentive - that the clinician that the outcome is attributed to will have some incentives to reduce readmissions. So, we were adopting a hospital one measure for clinician. We had some principles, we tried to stick to that manifests and sells in some of the details of the measure.

We did try to identify clinician and a socially groups who could positively influence the transition of care from hospital to outpatient setting or their decision to return and that was based on both looking at the data as well as looking at the research. We tried to align the measure as much as possible with the original hospital measure. So, you know, one for instance, we did not try to revisit a risk factor selection.

As Karen said these measures comprised of five distinct service line and cohort models that are estimated and rolled up. It is in reinvestigate whether that should be, you know, fixed for some different number we kept the same definitions. So we did this, you know, there are practical reasons to do it but also conceptually we wanted clinicians to feel like, you know, the measurement - that the measure was the same for hospital and clinician.

One important decision we made was that we rather than to adjust for hospital quality. We will assume that clinician performance reflected the quality of the hospitals where their patients were discharged. The – we, in looking at attribution, we assume that inpatient outcomes were most positively attributed to inpatient clinicians. That's consistent with literature, CMS policies, expert input we attained.

So there is also some thinking that primary care physicians, there is some evidence that primary care physicians do have influence over where the patients return to the hospital. And then, you know, we tried to consider the unintended consequences when we looked at clinicians that could be identified.

We wanted to make sure that there were incentives were for collaboration and not for anything else. So the measure we arrived at which you all have had a chance from yourself was driven - was developed under this principle. And we did engage throughout the process, not only a technical expert panel but also patient kind of work groups or some stakeholders who have contact with the hospital and we're familiar with the challenges of not returning the hospital. The evidence base we do, you know, Leslie was asking about that. There are some, there are a few but there are somewhat solid studies showing that there is a strong relationship between the admission of all care for the discharge summary and whether the patients returned to the hospital. This is under the control of the discharge clinician, which is one of the attributed clinician in our measure.

We found that there's a study that found, that shows that after high risk surgery and encounter with a PCP is strongly associated with lower risk of readmission. Patients who had the discharge process more thoroughly explained to them. This is a different study by the surgeon clinician had much lower risk of readmission.

And the number of studies have just kind of specific interventions that clinician is care for patients in the hospital can or take to reduce their - to improve their transition out of the hospital and reduce readmissions. So, you know, there are some of the, some measure we'd be happy to share, address that issue.

All of the other as Leslie talk about the, I'm sorry Leslie what is your other point, was it the performance gap?

Leslie Kelly Hall: The performance gap, there was, it didn't look like that was present in the submission that at least I could find and I'm not a measurement expert. I mean, and so there was a question about it, you know, is it, my question, is this our performance gap or is this simply a way to measure whether a new financial model is effective, more effective than the existing financial model?

(Jeff Haran): Okay, well.

Leslie Kelly Hall: Okay, because mid is the new financial relationship with a seeker service.

(Jeff Haran): Right. Well, you know, our existing measurements when they don't, you know, measure in data that didn't involve any financial incentive. I don't, as Karen said the range of risk adjusted readmission rates for clinicians was from 5% to more than 38% for groups within 7% to more than 25%.

Whether that reflects just variation due to financial arrangements, I don't know. But we haven't tried to look at that.

Leslie Kelly Hall: I guess my question was because that's those have been - that's overall different gap. And that was trying to understand is there a gap between this and this measurement and they all cause readmission measurement of trying to understand that gap? What's the problem we're trying to solve, rather than the overall readmission rate?

Karen Joynt Maddox: Hi, this is Karen, this is Karen Joynt (unintelligible).

Leslie Kelly Hall: So I think that, you know, the problem that we're trying to solve is to align, you know, as you mentioned, physicians that practice in the hospital was that there hospitalists or primary care physicians, you know, so share attribution of the outcome to incentivize them to do the sort of, undertake the proven techniques that just went through to try to reduce readmission rates among their patients.

> So I think, you know, you have to sort of, I think it - I think the logic that we're implying. But I'll state it more clearly is that there's a systemic role for hospitals, a system that provide care to engage in efforts to reduce readmissions, while patients are in the hospital under their acute care. But

there's also some specific physician directed actions, right, that are as unique to the individual physician patient encounter.

Both a discharge is just Jeff had described and in the transfer of care to the outpatient a successful transfer of care to the outpatient settings that are equally important for us to incentivize in the right direction. So I think that's what we're implying but I'll say it more correctly by using this measure.

And I think it's a program priority. I don't mean to speak for CMS here but it's my understanding that it's a program priority to provide that types of measures for, you know, specifically for hospitals, who might be discharging patient to be able to assess the quality of the care they're providing, right?

And the group that's relatively under measures compared to some other practitioners.

(Lisa Souter): Sorry, this (Lisa Souter), just also from core, just drilling down on Karen, and Jeff comments about the fact that, you know, we spoke we had a technical expert panel with patients and both acute care and post-acute care providers. And we heard a lot about that one need for shared attribution in order to incentivize coordination. And the specific individuals that are being attributed to in this measure, as being, you know, critical liaison between the inpatient and the outpatient setting even when they may or may not see themselves that way.

Those are the names that are getting communicated to post-acute care providers, such as a discharge and clinician is the contact person for the Home Healthcare Agency and things like that. And so there was a lot of conversation among the clinicians in the expert panel about how to identify the responsible individuals to improve the communication and coordination of care.

And so that's a little bit of granularity. I think it's helpful to understand that the attribution schemes were both evidence based. And sort of who we could document with seeing the patients and taking care of the patients, but also driven by sort of on the ground logistics that the expert clinicians and patients were bringing in front of us.

Wesley Fields: This is Wes. I've got a couple of questions for the developers. The first is since this is really an adaptation of a hospital messages that's used to benchmark all hospitals in the case of 3.95, you know, we were just talking about physician providers part of the payment. So it really is apples and/oranges and its financial models go.

But the question I have is, was there was the primary interest of the folks that are stakeholders and sponsors outside the CMS to just use this for a specialty level quality measure for MIPS or more the group level tax ID driven. The focus on something which potentially could be used for as the basis for an APM so is this just about specialty level MIPS measures or is it more about APMs.

John Bulger: APMs, I'm sorry.

(Lisa Souter): Those are Alternate Payment Model. Alternate Payment Model, yes, yes, okay.

(Lisa Souter): And the answer is no. The intention for this measure is it's used in MIPS.

Woman: I'm on the left side.

NATIONAL QUALITY FORUM Moderator: Benita Kornegay Henry 06-21-19/3:25 pm CT Confirmation #21926731 Page 28

- (Lisa Souter): I think the goal was to engage as many clinicians responsible for care of the patients as possible in the context of the MIPS program.
- Wesley Fields: All right. So a follow-up question then. The five subsets of the patients and services cross several medical specialties. So, if this is primarily about MIPS and we're looking at about quality benchmarks at the specialty level. Do you feel like you have equal involvement and engagement by neurology and internal medicine?

I mean, the obvious part of this is the hospitalist piece of it really primarily to internal medicine specialist and family and medical specialist. I'm just not clear about how engaged and enthusiastic folks from say neurology or cardiology are in the underlying five groups of services that better bundled up here.

John Bulger: You know, I think that what we heard from clinicians was that they were, you know, including neurologists as they were, you know, pleased to have a measure that they could - that they could use. I think there was a strong sense among other stakeholders that it made sense when patients, you know, who were in the hospital for, you know, logical care, you know, would have their outcomes attributed to someone.

I don't know, I guess I'm not quite sure where you're going with, you know, the five courts. It's true that, you know...

Wesley Fields: Well, I'll tell you where I'm going. I'll explain where I'm going. For the sake of some other folks in the work group, the reason this matters to me is that 1789 is a relatively level plain field.

In that I checked with some of the support documentation in CMS overnight. And the best I understand it, all facilities get benchmark around 1789. And it's a revenue neutral thing where their performance around this and many other measures in a compact way will dictate a level of reimbursement per facility services for party services.

On the other hand, the way I look at this and I'm inviting you to tell me I'm wrong, if this is really for Part B providers, even though I have strongest part of idea trying to get the better alignment around important outcomes that are evidence-based and that's, you know, that potential is very true about this.

You know, the differences here I think is the physician providers probably have more of this discretionary control around whether or not they choose to adopt this as a quality measure for themselves under the niche program. You know, that's the essence of macro.

So to me, there's a lot more optionality at the provider or the group level for how 3495 would be used in part as a financial model, if you will. Then there currently is for the existing hospital measure 1789. So, did I misstate that?

John Bulger: My understanding is the same as yours. They do have more flexibility and use this measure, you know, our physician developers is to try and construct a measure which...

Woman: (Nina Claire).

John Bulger: You know, which best measures performance of these providers, I don't know how to...

Wesley Fields: Well, let's just say this later in our discussion, when we get to the significant gap in outcomes for dual eligible, I think there's something like a 5% gap across the datasets for the rate of readmissions for dual eligible versus traditional beneficiaries.

And I guess, when we get to that point in the discussion about some of the technical aspects of this, I think that unlike hospitals that had to perform around 1789, I actually think that there's a significant arbitrage opportunity for groups to decide whether or not to use 3495.

That in part will be driven by the patient population they serve and their sort of grass roots level understanding of the relative risk. And I think, you know, since that also is revenue neutral in terms of how it moves the structure. It's actually not the same type of alignment you see with hospital measures or with 1789 in particular.

And, you know, I strongly support the direction of this measure. I just think it's important for folks to understand that whether or not physician groups choose to use it as part of their participation will likely be driven by their better understanding of their patient population then probably you folks did in developing the measure.

Lisa Souter: Thank you. This is Lisa Souter from the developer. As the developer addressing one of your earlier questions, I just wanted to follow-up that we did have a range of different clinical expertise. You raised cardiology and neurology, we had a cardiologist, we had several nurses, other care level care providers, and anesthesiologist.

> We did not have a neurologist on the technical expert panel but we did receive a comment from the American Association of Neurology that supported the

measure and asked for some additional kind of tweaking of the underlying cohort in public comments.

And then, I also just wanted to follow-up that, that this measure, actually, another version of this measure is already implemented in this. And because it doesn't require physicians to submit data for calculation, it's calculated for all physicians in the country that meet the volume and MIPS eligibility requirements.

That current version of this hospital wide readmission measure attributes the outcome to a primary care - they called the primary care provider, which is the individual who's seen the patient the most in the 12 months' measurement period.

And we are revising that measure to attribute not only to a clinician that we can document, has seen the patient pre prior to, I mean from a primary care standpoint, who've seen the patients prior to the admission and any readmission outcomes but also to reflect that the inpatient care has a lot to do with readmission risk.

So this is actually a re-specification of a measure that is in use in MIPS and calculated on all clinicians. I will say that you are correct. My understanding of all the rest of MIPS is that clinicians have a lot of flexibility in what they, in which measures they report on and I believe that still is the case.

But even in that situation, physicians are receiving information about their performance on this measure regardless of whether or not I think, is use in their calculations and I apologize that I, you know, I can't speak for exactly how CMS is intending to lead this measure going forward.

But because it does not require clinicians to submit data to calculate it, the information can be shared back with clinicians on a much more universal basis. So, there's much less of a risk of clinicians not receiving information about patients that they have participated in their care.

Wesley Fields: Well, I just want to flag one other thing. We can come back to in terms of the discussion when we get to the downstream parts of your analysis and the reviewer's analysis. But I think that the headline here is that whether it's signal to noise or other validation measures.

This looks like a better measure for groups than it is just for individual providers in terms of its power or its solidity. And I think that'll probably be borne out in the details from the reviewers.

- John Bulger: Great. Thanks Wesley. Does developers have anything else that they want to add before we go forward with the rest of them?
- Man: No, I don't think so.
- Lisa Souter: Yes, we can address questions as they come up.

Man: Yes.

- Lisa Souter: Thank you.
- John Bulger: Great. And (Wes) did you have anything else on important that you wanted to...

- Wesley Field: No, actually, I think the measure worksheet for this one is really helpful and to the best of my ability I'm going to try to stick with what's there from the reviewers.
- John Bulger: Thanks. Any questions from any other members of the Committee on importance?
- Leslie Kelly Hall: Can I just echo the comments about the group? This is Leslie, because if we're looking as the goal is shared attribution that is the goal for this. A highly functioning ACO will be able to care for any other individual physician patient and be able to coordinate care effectively.

So when we're attributing it to an individual submission it's somewhat diminishes the effect of a highly functioning ACO that is doing shared care, shared attribution, shared knowledge about a patient. So I just would like to add that to rest this comment.

John Bulger: Yes. Any other comments or and I just remind people if you're on the phone while everybody's on the phone but the mute if you're not talking because we've had some background noises come in with that.

And I would agree with what you just said. I mean, the essence of hospital medicine is it's a group, not an individual. So, you know, maybe the use of the - using a clinician who build the most through the hospital stay, you know, might get to that.

But I just think, you know, having spent 15 years running hospitals' program, we didn't attribute things like this to the individual. They were attributed to the group because you had multiple people coming in and out taking care of the patients and, you know, the discharge physician may, patient may I don't like to stay a five days and the discharge physician may have only been taking care of the patient the last day.

But they are going to be one of the people, they are going to have patients attributed to them using this measure. So I, you know, I think this notion of attribution is problematic in the measure. But I'm sure we'll keep getting to that as we go through.

So, Erin do we need to, we're going to do voting later right by, so we can just go through the different topics?

- Erin O'Rourke: That is correct. We did add the survey link to the voting tool in your invite if you...
- John Bulger: But we're not voting real time. So we can, you know, we're done with importance, we can go to what's next. And that's...

Erin O'Rourke: So importance with evidence and gap, so there's nothing else on either of those topics. We can move on to reliability under scientific acceptability.

- John Bulger: Okay. Let's do that. (Leslie), you want to start the reliability one?
- Leslie Kelly Hall: Okay. So, it looks like all the way around there. The group said there was moderate reliability and validity. The concerns that were raised within it were some of that we've already talked about the individual physician or hospitals or makeup. Also there was a question about the logic model used to support this look exactly the same as the 1789. So again it's the logic and the rationale behind this, how is it demonstrating a difference from 1789 other than the financial model? So I think we've beat that horse, discussed that already.

But again these go through, that papers the whole document. Also on the social risk factors and comments there was a discussion about the role, I think it's on Page 14. If we're not adjusting for social risk factors, the measure developer, it's a demonstrated adequate risk adjustment and to account for the medical factors.

And again, it seems to reference 1789 as the background of all of this and is there any further or additional information about either a logic model or the rationale associated with it? Other than that I think covered it. Thank you.

Erin O'Rourke: And this is Erin. I just want to jump in with a clarifying comment on Leslie's point. Seventeen eighty nine is currently endorsed at the facility and ACO level of analysis. This new measure would be seeking endorsement at the clinician individual or clinician group practice level of analysis.

So that's perhaps a key difference to highlight and to maybe address some of Leslie's questions. Also, this measure does use the multiple attribution model that's not currently or is not in 1789.

So there're some - a few differences that the committee may wish to consider there that I did want to point out for endorse measure.

- John Bulger: Yes. And Erin, while you're saying that, is there ability to split endorsements between individual and group or is it in mass when we end up considering this?
- Erin O'Rourke: That is a very good question. The developers submitted it for both. I, maybe I know Lisa's on the line. I don't know if they have, if the committee has the ability to endorse at one level and not at the other or how that decision works.

- Lisa Souter: Yes. So perhaps and sorry I joined late. So, and I joined some of the discussion late as well, they could recommend, you know, that it, you know, they have a preferred level. And I don't know what additional recommendations you want to go with the endorsement, their recommendation for endorsements decision.
- Erin O'Rourke: So maybe, while it's currently submitted for both and I believe there is testing for the individual and the group level. I think maybe based on the committee's conversation that is probably something we could work with the developer if it should only be resubmitted at the group level.

But I think right now, perhaps (unintelligible) to consider what's in front of you and to make the decision in that framework that they've applied for endorsement at both the individual and group levels of analysis.

John Bulger: So, we're going to consider it as an and not an and/or?

- Erin O'Rourke: Correct. Let's consider it an and for now and things that's been going so we could...
- John Bulger: Great. So great, Les or Karen, do you have any other comments on for liability?
- Karen Joynt: Yes, actually I do. There's a sort of an ecological feature of this measure and how it'd be used for payments under MIPS. I mean, I, you know, I'll just say that I think that the analysis suggests that the groups that are most likely to perform well under this on a voluntary basis are those that are, that do have scale, that do have better integration both with the community and the facility.

But in terms of how MIPS and Macro are structured, I just want to point out that the revenue neutrality is achieved because whatever the reward is under MIPS is essentially a takeaway from other providers in Part B who figured to be in smaller groups or in disadvantage to clinical settings, who either had trouble reporting or have trouble integrating.

And so there's a consequence here where because of Macro's design, the groups that perform at best under this kind of a model are sort of compounding the losses if you will, from providers that are less well-integrated or that are in communities that have less cognitive care post discharge. And, you know, that's the reality of the system.

But I actually think that's quite different than 1789 where I think you can argue that this is a more level playing field for all the facilities. And it's also embedded in a much more, you know, much larger basket of outcome measures.

John Bulger: Great. Karen, anything on to add?

Leslie Kelly: No, I'm asking in terms of their liability of the measure itself. It actually is a little bit higher for the clinician level than at the group level which I suppose seems a little bit counterintuitive. Except it maybe it's true that within clinicians there's just less variability in either patient population or sort of practice patterns and there is at the group level.

So I don't at least based on the reliability testing and it doesn't seem that the clinician level is less reliable. Now, both of them come out with the reliability of, you know, 0.54, 0.64 I guess, which are right around that sort of moderate cutoff for lack of a better term, even though we don't have cutoffs that we specifically go for, actually the group on like I said, being a little bit worse.

So it depends on whether you think that's adequate in terms of the technical reliability of the measure.

- John Bulger: Great. Any other committee members have any questions or comments from a reliability standpoint?
- Christie Travis: And this is Christie and my question is not so much reliability but just to help me frame my thought process here. If I am remembering correctly, this is like one of the few maybe the only measure we've looked at that actually specifies a particular program in its title.

And, you know, so I can appreciate the challenge, I'm trying to figure out, I mean, I guess I'm looking to NQF to give me some guidance on how to think about this in terms of how, you know, how the program operates and how this measures then sits in the program or are we still looking at this just for the measure itself, regardless of the program?

Karen Joynt: Sure, that is a great question. So, the NQF trust us, is, you know, continues to be use agnostic but I think here we do have the benefit of the developers and steward of being upfront that this measure is intended for use in the summit program.

And it's probably worth keeping in mind that it seems based on Lisa's comments, this measures would be intended to replace the current readmissions measure in the mix program. So, readmissions is one of the measures used under the quality domain of MIPS. And to maybe echo Lisa's comments that it's not a measure clinicians select. If you meet the eligible threshold, you would be assessed this measure.

So, I think it's perhaps important context that the committee may wish to consider that this measure has a very specific purpose in mind. But and obviously others wouldn't be precluded from using the measure, you know other payers is the things we have endorsed.

But it's perhaps a little bit of unique situation that is very transparent that we've to see intention of this measure would be and it would perhaps be to replace a measure that's currently in use in a specific program.

Jeff Hering This is Jeff Hering, if I could just jump in a minute. Yes, part of the use of the program in the title reflects the fact that this measure was designed to measure a certain set of clinicians. Not all clinicians, not just primary care clinicians, but also not clinicians who don't have, you know, patient-facing care.

And so, you know, we identified them as themselves well because that was, those are the set of clinicians we designed to measure. So, you know, that's one reason for the language there, I just want to say that.

(Susan): And this is (Susan) from the core team. I just – I wanted to see if we could speak a little bit more clarification. You know, measures often come through with different levels of potential use. And I was surprised to hear you say that the committee couldn't weigh in on one versus the other that that would require a resubmission, which seems like given that the information here has been provided the committee with all the information they would need at both the individual and the group level.

I just wanted to pause again, to make sure we understand correctly that this committee does not have the ability to put forward a preference that would be endorsed at one level versus the other. And that would actually require an entire repeat of this process with the same information coming in from the committee and is that what you're saying?

Karen Joynt: Sure, no. Thank you for the clarification. I think I didn't want the committee to not consider your application as submitted since you, as you know that you have everything for both levels of analysis. I think if the committee has a strong preference that perhaps they would want to only endorse at one level of analysis over the other. And that's the way the conversation goes, we could work with you, perhaps to withdraw the measure from this cycle and reconsider it in a future cycle.

> It's a little, again, I only search through and I don't have a precedent in mind for how the process works where the committee would maybe be willing to endorse at the group level, not at the individual. And if that's something they can change midstream like this, based on the committee's conversation or if we ask again to...

Karen Joynt: Actually this is what I have a sort of a point of order question. I think I know the answer but I think this is probably a good moment to ask a question.

I'm assuming that 3495 could be used by a group providing these kinds of hospital centric services while also participating as an ACO with a similar hospital wide readmission measure. And that all this could be nested within a healthcare system or, you know, a partnership with either one hospital or with the, you know, the system of hospitals.

I'm assuming, that that the architecture of this measure would allow it to be used for 1795 or 98 or whatever it is as well as 3495 as well as the parallel quality measures for ACOs. It would be helpful for me in terms of the question you just raised to know whether or not I'm correct about that. Karen Joynt: Sure I can at least my interpretation from what the developer has submitted is that the measures are designed to be as harmonized as possible so that it would essentially work at the multiple levels of analysis. As far as overlap between the programs, I don't know if that's really a point of consideration for this committee and how the different programs that this measure versus 1789 could be used in our work together and any areas of potential overlap.

I might ask if anyone from the Yale team did want to speak how this measure would work with 1789?

Karen Joynt: The question I really like to have answer excuse me, the question that I really think is worth answering is for clinicians and physicians under Part B or who are participating in ACO, CS Medicare ACO projects.

((Crosstalk))

John Bulger: This is John. But I mean, if you're in an ACO and, you know, eventually they'll all be at-risk. But if you're in at risk ACO, then MIPS is a moot point. If you're in a non at-risk ACO, the quality score in MIPS is actually the ACO's quality score, not your individual quality score.

> So it's a moot point there as well. So, you know, this measure could be applied and you would get a number but the number won't matter because if you're a clinician.

Lester Karen: Yes, that's helpful John because if the concept in terms of architecture is that this is part of a continuum, as far as a degrees of integration or alignment that actually makes me feel substantially better so that's very helpful John.

John Bulger: Any other comments or questions on reliability?

NATIONAL QUALITY FORUM Moderator: Benita Kornegay Henry 06-21-19/3:25 pm CT Confirmation #21926731 Page 42

(Shareen): John, this is (Shareen). Before we move on from the question, I feel like there was still an outstanding question on the level of analysis. So, I mean, the only thing I would provide is that if the developer, the level of analysis is part of the specifications so the developer does need to provide whether it's at the group or the individual in this case but the developer stated at both.

And really the question from the committee is whether the testing supports both levels of analysis, if only one of the levels of analysis is supported by the testing. Then the Committee can determine that and only endorse the measure at either the group or the individual level of analysis. So that's really the way that this Committee is reviewed or, you know, considered the level of analysis as part of the specifications before. So it's not necessarily that the developer needs to resubmit the measure only at the group or only at the individual level.

But really the question is whether the testing supports, both as the developer is submitted in the current application in front of the Committee. So is that clear in terms of the process and President that the Committee's used in the past? And at least that there's other comments that you have on the sort of the point of order here, we can certainly clarify that. So I make sure that's cleared before we move on.

Yale team does that answer or address your question?

Man: I guess I heard two things. One was that the President was to endorse only at the levels that seems appropriate. And I also heard that we had two but, I'm still not quite sure what the President is.

(Shareen): Okay. So, the appropriate in this question is really what the testing supports. And that's essentially the termination that the Standing Committee needs to weigh in on that's the question. So I'm not saying yes, one or both is really what the testing support based on what the developers submitted.

Erin O'Rourke: Sure. And this is Erin, maybe I can clarify my comments. The developer did submit before with the level of analysis that they're seeking endorsement for as clinician individual and clinician group practice.

They submitted testing, demonstrating, reliability, volatility of both levels of analysis. I think if the committee perhaps wants to disentangle this and Lisa please correct me if this is, if I'm misspeakingly, we could try, we could take separate votes on with the committee supported at individual with the committee recommended at as a group practice level.

If that ceases the parts of question if the Committee feels there's material difference and that they want to discuss and vote separately on individual and group that might be the clearest path forward. And maybe I'll just pause for because I think Lisa is chiming in.

Lisa Souter: Yes, sorry. I was just going to agree with you on the two pathways. If there's not a comfort level or consensus amongst the Committee on the two different levels, it might be cleaner to do it that way.

Keith Lind: Erin, this is Keith. I just wanted to practically ask whether the SurveyMonkey will have options for voting twice on each of the elements, I have a...

Erin O'Rourke: Accurately set up, it is not so I think if the Committee feels you would wish to consider the two levels of analyses separately, we should probably pause on anyone using the SurveyMonkey and we can set up a new tool and send that back out.

NATIONAL QUALITY FORUM Moderator: Benita Kornegay Henry 06-21-19/3:25 pm CT Confirmation #21926731 Page 44

## Keith Lind: Thank you.

William Fields: This is Wes. I just what, I just want to reference on Page 7, bottom of Page 7, something for one of the reviewers which I think is pertinent to this and it's a quote near the bottom of that paragraph.

While there are many possible exceptions to this attribution model, it was the one tested with moderate to substantially reliable results. And I guess my point about this is an extension of John's. My group of which I'm in a (Meredith Partners), somewhere between several hundred thousand and a million hospital encounters a year.

And I talked to the folks that run the hospital program and the post-acute programs and I just want to underline what John said. It's, you know, encouraging that they got a moderate level of reliability from what they proposed with those three providers.

But I see most of those three categories as strong men or women for a much more nuanced and complex, the care delivery model that's going to vary by facility and by region, et cetera, et cetera. So, and we can maybe later, if we have time or if we have a need, we can talk about some of those use cases where, even though this may have moderate reliability, it's arbitrarily narrow in terms of trying to define what the measure aspires to do.

So, you know, I honestly think that the folks that use these are going to use it at the group level. And they're going to be careful with the fact that whether or not the actual providers are, you know, fit to form those three definitions are not will mattered less than they had a higher level of competence in the quality of the care, the outcomes they generate. And it's demonstrated their willingness to use 3495 of this. So, yes, actually it would be helpful. I think that SurveyMonkey reflected that because it's almost immediately. Although I understand that you're going to be collecting claims data profile every Part B provider in the country in theory.

I think that the reality is that this is because of the attribution problems and because of the variability in terms of care delivery, this really only has functional significance to me as a group doing it and the interaction between groups and facilities.

Lisa Souter: I would echo that as a patient as well because you want typically that your care is coordinated that the best example I've personally had in my family of great coordinating to health care meant the physician was walking in the hospital and to care for my family member was not as material as the information they had on hand as the standard of care they adopted as the level of care coordination they had amongst their group.

And the excellence that was demonstrated was by the group, regardless and independent of the individual physician that presented. And I think that's the end goal for all patient is to be able to have a confidence in excellence that is based upon quality across any member of a care team that's caring for me.

Karen Joynt: This is Karen. I would actually push back on those comments. But I certainly agree with the idea that it's important to have a group taking care of you but the reality of how outpatient practice is set up nationwide is that there are scenarios in which there's a sole clinician or a small group. There are scenarios in which there's a large group and the way that this measure gets used.

I don't know that we can necessarily predict. I think what we probably ought to be...

Lisa Souter: If it's a MIPS and ACO, it's not a small group.

Karen Joynt: No, it's everyone. People are still required to be in MIPS if there are small practices if they have adequate Medicare billing or attributes of patients in a year. So it actually can by now be pretty small group. And that's very separate from ACOs which again have their own readmission measure, which is not the one that we're talking about today.

> So, I think the question before us is more do we think that this can measure the quality of these clinicians. Because at least my understanding of the way that MIPS are being used is that there are a number of clinicians participating in the program as either solo or very small group practitioners.

> And anyone can actually elect to even if you're part of a large group. So, some of the details of how this program works especially because MIPS is likely in well, it's obviously in flux as it evolves over time. I'm not sure that we can sort of perfectly predict going forward.

Erin O'Rourke: So this is Erin. I did want to just echo Karen's comments there. Seventeen eighty nine is the measure that's used for ACOs and in the ACOs program. This measure is being put forward as a measure for the MIPS program. So, clinicians would not get to pick this. It would be the six measures you picked and then if you hit the threshold, you would be assessed by the readmissions measures.

Again I think I just wanted to disentangle that this measure would not be used in the ACO program. Obviously the developer highlights that they tried to synchronize their methodology but that this measure is for a clinician level program.

John Bulger: Yes, that's absolutely correct. Any other questions, comments on reliability?

- Karen Joynt: Just a reminder that it currently has a pretty high volume threshold so that how CMS visit in the future could change. But that impacts concerns about smaller groups. At this point, I think requires, I don't know the exact number but quite a large number of patients to be equivalent.
- Man: Two hundred.
- John Bulger: Yes, which most physicians may get a primary care level, you know, that which is the third option are not going to get to at an individual's clinician level. At a group level absolutely, there are, you know, smaller groups that would get to that but most care physicians are not going to have 200 admissions in the course of a year.
- Karen Joynt: Among the over 65 fee for service Medicare.

John Bulger: Yes, exactly. It's just that. I mean, that would be, you know, that would be someone that if you are monitoring that person, you would probably go look and see what they're doing.

So any other comments or questions along reliability? All right, let's move to validity and, Wes, you want to start there?

William Fields: Yes, I was going to leave myself on mute for a while. But sure, let me see if I can get there. Well, one thing I'd say in passing just to get myself teed up

here. This bit with, this is an old topic but the measure really restricts itself to inpatients of encounters.

So that there's obviously a lot of potential readmissions that are going to fall into the category of an ED visit or an extended ED encounter or an observations day. That is what it is. I just wanted to mention it in past. Let's see here. So with validity now, John, is that where you want me to go?

John Bulger: Yes, that's correct. Let me just ask one more question now that it came to my head with that last comment. So the current measure that this is replacing, I think what the people you all said essentially is the, you know, Option C, which is the physician who takes care of the preponderance of the care on the outpatient side, correct?

> And this measure then adds the discharging physician and the physician that takes as the preponderance of care on the inpatient side of the measure. Is that an accurate way to look at this?

(Susan Abronson): We also made a change to how that outpatient provider is selected. So, the current measure just looked at the preponderance of care in a year, the measurement year without regards for when the admission, the index admission occurred.

We changed that to look only one 365 days before the admission to ensure that the person attributed the outpatient provider had in fact provided the front and a care prior to the admission.

John Bulger: Yes. So, what this is really, you know, there's that small change and as you said down in the preamble, right now, there's not a whole lot of measures for hospital-based physicians, specifically hospital wise. So this change it has

that and there's that other measure, the one other thing that came to my head is that other measure NQF endorse?

Erin O'Rourke: Sure. This is Erin. We do not have NQF endorse readmissions measure at the clinician level of analysis. So the current version using the program is not endorsed.

- John Bulger: Okay, great. So this is kind of, this is adding two clinicians. And it's asking for endorsement, which doesn't exist right now for the primary care one and changing the methodology a little bit.
- Erin O'Rourke: And this is Erin. I did want to, sorry to jump in. I think Susan I apologize with you on the spot. I know the other measure may not be yours. But I think there might be some changes with the threshold. So and who would be potentially assessed under this measure. But I don't want to make you speak for things but I think may use it, CMS certainly.
- Man: We are not probably familiar with the existing measure but it does have a volume threshold of I think 200. It does attribute only two groups and as to benefitted attributes to whomever filed for the most primary care claims during the calendar year, which could be, you know, completely after discharge.
- John Bulger: Yes. Okay, great. All right, Wes ready?
- William Fields: Yes, sure. So, I'm just going to work down from the bottom half of 10,
   empirical validity testing, face validity. One of the respondents or actually a
   majority of the respondents seem to have some concern about attributions.
   The quote here at the bottom which in among those who disagree the primary

concern was the factors which led to the increased risk of readmission were beyond the control of any single eligible clinician or clinical group.

The counterfactual for that which to me is sort of more interesting and why I think there is an arbitrage aspect to which groups adopt this for MIPS is towards the bottom of Page 11, third in the last bullet. The developer then examine the strength and significance of the SDF's variables in each of the client specialty cohorts in the multi-variable models and found a modest of effect size.

Developer found that the addition of these variables to the model had little to no effect on eligible clinicians or groups. My understanding of that and I welcome to be corrected but I think the phenomenon here is something we've seen with SDFs and other categories, which is that a lot of this relative risk or need for readjustment if it exists, either at the facility or the community level and I think that it's highly likely that the groups, you know, know where they practice and know their patients they serve. So that's sort of the basis for one of their headline, you know, comments I made when we get into this.

The last bullet in this section at the top of 12 I guess it is or 11. The S&P said group members raise some concerns of social risk factors were excluded from the risk model given the effect and size, the potential for negative consequences.

The S&P panel members suggested the developer examine other clinical variables that could underlie disparities such as frailty or functional status. You know, I think, for anybody on the committee that this actually worked anywhere around that interface between the community of the hospital.

You know, I think there's a growing body of evidence that suggests that while there's a higher correlation within the first week or 10 days or possibly two weeks post-discharge. When you get up to 30 days, you know, this is like ice cream, it begins to melt.

And I think that the pertinence here is speaks to something we talked about earlier is that, you know, if the folks in the hospital are well aligned with the folks providing care and community settings, the post acute services and even in (unintelligible) type settings, you know, I think you're likely to see a better continuity of care, you're likely to see the aspiration goals made in terms of contacts with the primary care docs in their equivalence.

But, you know, in terms of how this is structured, the relative risk or variability is still sort of embedded in the measure and I really do feel like that's one of the things which makes the application of the hospital readmission rate somewhat different for meds than it is say for all Medicare participating hospitals. I'll stop for a second, because I'm not really trying to provoke folks, I just wanting to make sure that I'm not sort of misinterpreting or over editing the comments to reviewers.

John Bulger: Okay, great. Leslie or Karen? Karen or Leslie, whichever way you want to.

Leslie Kelly: Sure. I have nothing to add - this is Leslie. I have nothing to add. It still goes about attribution as the common theme.

Karen Joynt: So this is Karen. I'd add two things, I think both of which (Les) brought up. One is that I don't know that we knew this when we started looking at the admissions and readmissions measures and we just didn't think of it, but as you think that moving forward probably some understanding of the impacts of the observation phase is probably going to be really important to look at. One thing I didn't know until quite recently is just how different the use of observation states is across hospitals, because of differential facilities and sort of what you have and what you can offer and how that plays into whether or not you can transfer people or keep them on knobs and all that. And that's - that was all news to me actually when I started taking transfer calls as the overnight cardiologist from places that can (unintelligible).

So while I don't know that it should - it validate this, but I think that it may be something we want to keep an eye out for in terms of thinking about how differential use or differential access to or something like that that observation stays for both numerator and denominator may impact how this measure reflects real sort of revisits after a hospitalization.

I would also - just really surprised to nobody bring up the dual eligible thing. The rate is obviously markedly higher for duals compared to non-duals about five absolute percentage points in this investigation which is higher than some of the other places. We've seen although that's part of the event right here is higher.

The effect side was still modest even after controlling for other medical core morbidities although it dropped to be in odd ratio of more than 1.07 to 1.16 range.

And then to my earlier misguided comment, it is now relevant which is on page 70 of this document, the developers helpfully outlined the way that things change when adding the dual eligible indicators. So the overall changes quite small. You know, in the 100s of percent range which kind of makes sense sort of mathematically. But the range in the movement is it looks like negative 1.1% to 4% and for a clinician group if I'm reading this right negative 0.45% to positive 2%. And so on the base of, you know, 15% rates that we see, those are pretty big movements.

So to the prior comment about what do we think this does in terms of overall versus moving certain practices, I agree that overall it does very little and would argue that taking any variable out of the model would probably not do very much overall. That's not true that the standard we hold variables too in general. But there are groups on the tails for which this does make a difference in terms of the adjustments.

Man 1: (John), there is one thing I'll confess about validity that I didn't get to in the data dictionary. But in terms of exclusions from both the numerator denominator, things that are - they have a primary psychiatric diagnosis are excluded. The thing I'd like to feel better about is that things like delirium or other sort of organic conditions, especially in the elderly or those with preexisting dementia that they presented a change in mental status. I feel better about this measure if I knew that they didn't get dropped as being psychiatric disorders, maybe the developers can speak to that.

John Bulger: Any callers from the developers?

- Woman: This is (Susan). But the goal is to drop if it seems to be the primary reason for admission. The hope is that if it's something they developed like the lyrium during a hospitalization, those patients would still be included. I can't promise that it happens perfectly in this measure, but that's the way it's designed.
- Man 1: Okay. I'll accept that for now. It's a tough period. It's a moving target, even in training, you know, hospital physicians, the emergency physicians. I'm to page 14, like I said, so really the same comment here. I too like Karen was

sufficiently. I think that's probably all I've got for you, John. John Bulger: Okay, great. Thanks (Les). Any other comments from the committee around validity? Okay, I don't have this stuff in front of me, but I think feasibility is next, is that true Karen? Karen Joynt: That's right. John Bulger: All right. So Karen, you want to speak to feasibility please? Karen Joynt: Okay. This one was quite feasible, it's like it has been mentioned, it's being used in some form of claims collected and it already has a plan for implementation. John Bulger: Yes. Anything they had (Les)? Man 1: No. John Bulger: Okay, all right. So then use and usability, Leslie, you want to start that one out? Leslie Kelly: Yes. There was some discussion about how this was vetted and wanted to hear more about that. In the document there was a question about that is specifically in organizations that we're participating as specialty based organization. And then also I did have a question on the unintended

consequences under usability. When using individuals versus groups, is there

a potential negative consequence of inappropriately directing to individuals

struck by, you know, 5% variation in readmission rates that I mentioned

earlier for the dual eligibles. That makes me feel like a lot of this really is

about geography and markets, but the point I guess I probably would make

when a group member could adequately respond. So those - the second is my question and the first is responding to some of the comments in the document about vetting.

- Man 1: Yes. Any just before the developer just in case well, go ahead. Developers have comments on.
- Man 3: Sorry, I was on mute. I was a few times now. The on usability regarding I guess you're talking about sort of adverse selection effects.
- Leslie Kelly: That was the question my question, adverse selection effects and then the other was just generally there were questions in the by the panel about vetting.

Man 3: Vetting. I'm trying to find the questions about vetting.

- Leslie Kelly: How this is vetted under burden for and the risk and benefits, did the interventions match the objective. Those are some of the questions. I think it was page 22 was vetting how is the vetting done for this?
- Man 3: Okay. Well, I can't speak you know, we did not do a benefit analysis as we discussed this measure was developed for use in a program to replace a measure that already existed. So we did not have you know, in our public comment we did not hear anything about the burden of the measure, you know, as we mentioned before it's based on data that's already collected, the claims data, so there is no burden in the reporting. So I think from that sense, you know, it was not something that we focused on in our development.

Leslie Kelly: Fair enough.

- Man 3: In terms of adverse selection, you know, I think that is that is something we would be more concerned about and sort of a procedure based measure?
  Here, you know, I guess one can imagine scenarios where clinicians are you know, try to select, you know, patients who, you know, get admitted, because they think they'll have less chance of being readmitted. But I think...
- John Bulger: This is...
- Leslie Kelly: I my question is more about selection of doctor versus patient...
- ((Crosstalk))
- Man 3: I see, okay.
- Leslie Kelly: Is there a potential for delay in care, is there a potential for inappropriate response, because we're measuring at an individual level rather than a group level, especially in the inpatient study.
- John Bulger: Yes. I mean here is an example will be is I'm I mean I think this get both ways, but I'm a hospitalist and I have a patient that I'm pretty sure is going to readmitted, because they're difficult for whatever reason. So I could discharge him today or I'm going on service, I can let my partner discharge him tomorrow. At a group level it doesn't matter, because it's the group, but that's one place. And then from a patient standpoint, you know, patients trying to select, they don't really have a selection availability with these when they go to the hospitalists for example. So, you know, just get who comes, so what does it?

Karen Joynt: So this is Karen again from - you know, I just want to, you know, point out that when measures are relatively new even those measures already in the

program there is not, you know, any real length of experience with the measure. Although these general concerns about unintended consequences are vetted with our technical expert panel, that's certainly one of the things that comes up in conversation and is dealt with in the development of the specifications and in the testing. We just don't - you know, because the measure is new, we just don't have a lot of experience with its impact, right in the real world.

And so, you know, typically conversations beyond what made a reason in stakeholder today during the measure development conversation is about actual monitoring of data to support unintended consequences tend to happen more with endorsement maintenance, right, when the measure has been in the world with (unintelligible) and we develop some facts of experience that we can look at in the data.

((Crosstalk))

Man 3:	I'm sorry go ahead.
((Crosstalk))	
John Bulger:	Does that answer your question?
Leslie Kelly:	It does, thank you.
John Bulger:	Okay, go ahead (Les).
Man 1:	Yes, I'm sorry. I lost back and I was out there for a sec. I just want to get the developers a couple of use cases. Just to sort of expand, you know, (John's) concern and make it also a little bit less personal or provider specific, you

know, you've got five large cohorts of services and the (Med Surge) thing, I just want to point out that's reality in most settings large or small is that somebody that's admitted for elective surgery say joint replacement or something that's more to do with cardiology. It's highly likely that the interventionalist is not going to be either the admitting doctor or the discharge clinician.

And so that's one large example of hospital admissions where this - the three providers that this measure captures not necessarily going to be the one that has the most impact on outcome or risk of readmission. The same thing is generally true these days about stroke care. The key assessments likely to be, but you're all to stop and tell a neurologist and although you're - so you've got large numbers of neurological diagnosis that are another of your five cohorts. In many cases, you know, that assessment or that intervention is, you know, not something that the hospitalist or the post-acute folks have really much control over.

So I just - it's part of like practically speaking testing aside. I think this measure probably will work better with at the group level than at the clinician level, because when it comes, there is a large number of hospital inpatient counters in terms of categories and it not only touches lots of patients, it touches many other provider categories that actually aren't measured.

John Bulger: Thanks (Les). Karen, you have - and I'm...

Man 3: Can I respond to that or should I wait?

John Bulger: You can. If it's different from the previous response regarding unintended consequences...

NATIONAL QUALITY FORUM Moderator: Benita Kornegay Henry 06-21-19/3:25 pm CT Confirmation #21926731 Page 59

## ((Crosstalk))

- Man 3: Yes. I just wanted to correct part of the statement which was that the you know, we do attribute to the primary care physician and to a discharging condition, but not to admitting physician. The third attribution is to the clinician who builds the most charges during the inpatient stay and we chose that attribution, that is it came out of our stakeholder discussions, because it typically identifies say the surgeon or the interventionalist who does sort of provides a significant care during the stay. So I just want to correct it. It's not the admitting clinician, but actually...
- Man 1: Yes. Well, you know, we're running out of time, but I would challenge that and I'll give you a counterfactual. If you're looking at the plurality of encounters before the admission, for the very fragile patients that lived in or near skilled nursing or long term care facilities, it's highly likely that the plurality of claims are going to be somebody who is seeing them in a chronic care or a post-acute care setting...

((Crosstalk))

John Bulger: But I think it's the question - I think it's during the admission, not to interrupt, but (unintelligible) people said...

((Crosstalk))

John Bulger:...so the three categories or the highest build during the inpatient stay, the<br/>discharging physician and then the outpatient physician has the highest...

((Crosstalk))

Man 1: All right. I'll sign off from that, my apologies.

John Bulger: Karen, do you have anything quick to add on from a validity or excuse me, use and usability standpoint?

- Karen Joynt: No, thank you.
- John Bulger: Okay, great, thanks. So and that's (Aaron), right, that's it for the discussion standpoint.

Woman: Yes. Did you absolutely need to open for public comment though? Before we do - I open the line...

- John Bulger: Yes. And the one thing I wanted to say, I mean I think unless people have a strong objection it sounds like from the discussion from a voting standpoint, it will be helpful to be able to vote with either or not and.
- Woman: Yes, that sounds good. So we will update the survey to give those sorry, I cut someone off.
- Leslie Kelly: This is Leslie. And just related competing measures, just because this group had so much confusion about all of the different measures competing who participates in zone, I just would encourage that the developer in the narrative that describes this also describes the relating measures and how they've been harmonized a little bit better.
- Woman: Sure, thanks Leslie. And we can have a related and competing conversation when we get to the post content call. We can see what happens with the committee's recommendation on this measure and then once we know an endorsement decision we can discuss related in competing, because I think

that's a good point that there is some confusion. It'll be helpful to hear how the measures are harmonized and how they work together.

I did want just a couple of voting things. So we'll follow up with the SurveyMonkey that gives you two choices. We'll ask you to make a recommendation for the individual clinician level of analysis as well as the group practice level of analysis. So please hold off on voting and we'll clear the votes and get you a new survey link.

And then finally I did just want to - one process point, the decision in front of you is really to endorse or not endorse this measure on its own merits, not to make a decision about using this versus a new - reverse is the current methodology and that's more of the domains of the measure application partnership. So just a reminder of the focus of this committee, it's an endorsement decision on the merits of the measure in front of you.

John Bulger: Great, thanks, (Aaron). Well, can we open up for public comment?

Woman: Hi, good afternoon. My name is (unintelligible). I'm with the American College of Cardiology. I just want to thank the committee and the developers for the discussion today.

> I just want to echo a few points that were discussed today. One is that attributing this measure at the clinician level would result in small sample sizes that will be subject to large swings and performance and low levels of reliability and validity.

And I think the developers may want to consider raising the case minimum to ensure higher reliability of the measure. Also, it will be great for attribution to be at the 10 level with the requisite and first dispersed statistical reliability. Also, the diagnostic categories for cardio respiratory and cardiovascular probably needs some tweaking and refinement. For example, TAVR which is Transcatheter Aortic Valve Replacement is an increasingly large driver of valve related disorders for which there is a defined therapy and lower readmission rate. So (unintelligible) this procedure with heart valve disorders would be a mistake and the category then becomes too heterogeneous.

Also, it's not already left ventricular is this device - devices home no renown therapy and transplant should all be excluded. And finally, I don't know how much social determinants of health were informed in the measure development. But I think those should be factoring in the calculation such as transportation access and ability to afford medications and that kind of thing. So thank you.

John Bulger: Great, thank you.

Koryn Rubin: Hi. This is Koryn Rubin from the American Medical Association.

- John Bulger: sure, go ahead.
- Koryn Rubin: I appreciate the yes, we appreciate the opportunity to comment and for the engaging conversation today. The AMA is concerned with the weak evidence to support application of this measure across multiple physician in group. This measure would (unintelligible) broadly to any physician or group who meets the minimum requirements and under the NQF testing in six years it will be 25 patients. And this is so to clarify, this is a claims safe measure and under MIPS you do not have a choice for the all cause readmission measure. CMS would automatically attribute the measure to either the individual or group as long as they meet the case minimum.

In addition, under MIPS hospital based physicians including hospitalists have the option to have the hospital value based purchasing score applied to them. So that would mean that Measure 1789 would still apply to many hospital physicians and not this new measure.

And also in addition, physicians that practice - physician practices that participated in the shared savings program are held to 1789 and not this new measure. 1789 has specifically been tested and endorsed at the ACL level and they're not considered part of MIPS. So MIPS physicians are those that are not in shared savings programs or MIPS APM or an APM. Therefore that AMA believes the attribution must be determined based on evidence that the accountable unit is able to influence the outcome which also aligns with NQFs most recent improving after infusion models report.

This principle is also aligned with the evidence requirements for outcome measures in the NQF measure evaluation criteria which requires that there'll be at least once dropped structure or process that can influence the outcome and this relationship must be demonstrated through empirical evidence. Therefore, we request Yale or CMS begin to begin to demonstrate that these relationships with the accountable unit prior to implementing this measurement and we don't believe that this has been adequately demonstrated.

We agree that there is evidence to demonstrate that improve care coordination and programs focused on discharge planning can lead to reductions and hospital readmission. I thought most of the sighted evidence involve included multiple partners and clinicians such in health system, hospital nurse and/or pharmacists. Therefore we don't believe there is sufficient evidence that was provided to support that an individual physician or practice using the proposed attribution approach in the absence of some coordinated program or targeted intervention led by the health system or hospital can implement structures or processes that can lead to improve outcomes for these patients.

Furthermore, continue of care requires smooth transition to prepare patients changing clinical and social needs. However, the Stark Law often impedes the continuous and care transition. Specifically, in certain circumstances, physicians are prohibited from employing care coordination strategies on behalf of their patients. For example, an arrangement that pays for a nurse coordinator supporting a recently discharged patient care among the hospital, physician specialist or primary care physician due to concerns that this may induce future referrals to their own office to avoid an unnecessary readmission to the hospital.

As a result, we do not believe assignment of responsibility of the reduction of readmissions to multiple physicians and practices in MIPS is appropriate nor as the developer provided sufficient in termination to support the attribution of this measure to up to three physicians or practices.

We're also concerned with the reliability for based on the minimum case number of 25 patients. We believe that measure must meet minimum acceptable threshold of 0.7 for reliability and during the Yale public comment period in December of 2018 the data that was provided at that time was using at least 100 patients and yielded a mean signal to noise results of 0.991 for eligible clinicians and 0.997 for eligible groups. Therefore we requested the Standing Committee evaluate whether the case minimum of 25 patients is acceptable given the low reliability results.

We're also concerned troubled to see that no evidence or testing has been provided to support the attribution of this measure to the three distinct groups: discharge physician, primary inpatient care provider and outpatient primary care provider. While correlation to the hospital's overall star ratings and readmissions (unintelligible) and the star ratings are useful, we also believe that the developer has provided sufficient information as it relates to the measures application to each of the accountable unit to which the measure is attributed.

In addition, the conceptual basis used to explain why social risk factors were tested in Section 2B338 solely focused on the hospital and was not specific to physicians or practices. Therefore, it's difficult to determine whether additional factors should be considered.

Given the measure specifically developed permits, so a developer must also perform testing that demonstrates how the measure would perform under the MIPS benchmark methodology and physician compare star rating since CMS uses - utilizes two different methodologies for ranking and profiling physicians under MIPS.

In conclusion, (unintelligible) desire to apply these measures to the broadest number of clinicians as possible with the unintended consequences of inappropriately attributing measures to physicians for which they cannot meaningfully influence patient outcomes. The AMA requested the Standing Committee to carefully consider the potential MIPS information that could be provided to patients and caregivers if the measures do not have the clear evidence based to support attribution of the outcome to specific physician and can potentially produce for that are invalid and unreliable. Thank you.

John Bulger: Thank you. Other members that public that wish to comment? Okay, hearing none. (Aaron), do you want to speak the next step?

Woman: Sure. So next step, we will create a new SurveyMonkey and send that out, so please hold off on your voting. Otherwise, thank you everyone. We recovered the agenda items we needed to. So we will be cancelling the call for the 27th. We will be in touch about rescheduling - the request for reconsideration from last cycle for measures 3443 and 3445 at a time where we can ensure we have quorum of the committee. So for next step, please check your email for scheduling for that as well as a new link for voting for the survey. And we will be back in touch with the results once we have them.

- John Bulger: Great, thank you. Thank you everybody. Thanks everyone for hanging in there. Sorry, we were...
- (Group): Thank you.
- Woman: And thank you (John) and (Christi) for your leadership.
- Woman: Sure, bye.
- John Bulger: Thanks (Aaron) and team, bye-bye.

Woman: Bye.

END