NATIONAL QUALITY FORUM

Moderator: Endocrine Standing Committee
January 22, 2015
2:00 p.m. ET

Katie Streeter: Hi, good afternoon everybody. Welcome to the Endocrine Cycle 3 Measure Review and Evaluation Webinar. This is Katie Streeter, Senior Project Manager here at NQF. I'm joined by Kaitlynn Robinson, who is our project analyst and we also have Karen Johnson, who is a senior director of this project.

I also have here with me Ann Hammersmith, who is our general counsel at NQF and she's going to be opening up the call with roll call and talking about disclosure of interest.

Ann Hammersmith: Hi everyone. This is Ann Hammersmith, Interest General Counsel. And as Katie said, we'll be combining introductions with the disclosure of the interest. If you recall from last year what the process was where we went around a virtue of payable and disclosed. We'll be doing that again.

You recently received a form from us where we asked you about your professional activities and asked you to disclose anything that you thought was relevant. What we're going to be doing today is I will call your name and you can tell us if you have anything to disclose.

We don't want you to summarize your resume. We don't – we like you to disclose things that are relevant to the committee's work. Just because you disclosed does not necessarily mean that you have a conflict of interest. Part of the reason, we do this is in the spirit of transparency so that everyone knows where everyone else is coming from.

We are especially interested in research grants or consulting that you may have done but only if it is relevant to the committee's work. I also want to remind you that you sit on the committee as an individual. You don't represent your employer. You don't represent a professional society which you may be associated. You don't represent anybody who may have nominated you to be on the committee. You're on the committee because you're an expert and we're interested in your individual opinions as an expert in this area.

So, with that, I will call the names and tell us who you are, who you're with and if you have anything that you would like to disclose. I'll start with the Chair, William Golden.

William Golden: Yes, hello, Ann. I work as a medical director for Medicaid program, in addition being at universities so I do use these measures or use measures to design the incentive programs and I've also sat on the executive committee of the Physician Consortium for Performance Improvement but I have not been sitting on anything related to endocrinology recently.

Ann Hammersmith: OK, thank you. James Rosenzweig? Is James Rosenzweig on the phone?

OK, Robert Bailey?

Robert Bailey: Hi, I'm Bob Bailey. I work for the health economics and outcomes research team at Janssen Scientific Affairs and in my capacity, I developed evidence to support the pharmaceutical portfolio for Janssen Pharmaceuticals and I'm a stockholder with Johnson & Johnson.

Ann Hammersmith: OK, thank you. Tracey Breen?

Tracey Breen: Hi. This is Tracey Breen. I have a new role since our last meeting. So, I'm now chief medical officer at Mt. Sinai Roosevelt Hospital. I have no conflicts of interest regarding the development of these measures other than being listed in them as – hospital. Thank you.

Ann Hammersmith: OK, thanks. William Curry?

William Curry: Hi. I'm the family physician and professor of Family Community Medicine and Public Health Sciences at Penn State University's College of Medicine in Hershey. I have nothing to disclose.

Ann Hammersmith: Thanks. Vicky Ducworth?

Vicky Ducworth: Hi, I'm with the Boeing Company in our health policy and strategy and I'm with our clinical programs and delivery system innovation. I use some of these measures as an incentive within our ACO contracts for our own panels.

I have nothing to disclose regarding for the patient any outside activities related to endocrine.

Ann Hammersmith: OK, thank you. James Dudl?

James Dudl: Hi, I'm with Kaiser Permanente National Leads for Diabetes. I use these measurements for quality improvement and nothing to disclose.

Ann Hammersmith: Thanks. Ingrid Duva? Is Ingrid Duva on the line? Starlin Haydon-Greatting? Is Starlin Haydon-Greatting on the line?

Starlin Haydon-Greatting: Yes.

Ann Hammersmith: OK.

Starlin Haydon-Greatting: I'm director of clinical programs for the Illinois Pharmacists Association. And I'm a clinical pharmacist and that majors in pharmacoepidemiology. I serve on a pharmacy called The Alliance and the measure development that's being measures. We are developing our – in primary drug, (Adaferin) and so I have nothing to disclose.

Ann Hammersmith: OK, thank you. If I could just ask people on the line, if you were not speaking, if you can mute your line because we tend to be getting some feedback. Thank you. Ann Kearns?

Ann Kearns: Yes, this is Ann Kearns. I'm an endocrinologist at the Mayo Clinic and I have nothing to disclose.

Ann Hammersmith: Thank you. Sue Kirkman?

Sue Kirkman: Hi, it's Sue Kirkman. I'm at University of North Carolina. A couple of disclosures, I do clinical research and I do have some research funding that goes to my university for clinical studies from Novo Nordisk, which is a company that makes diabetes medications.

And, that's my only sort of financial conflict of interest. I also – from 2007 to the end of 2012, I worked for the American Diabetes Association and as part of that period of time, I did some work with NCQA, you know, who of course is a measure developer as part of their diabetes expert panel or expert advisory panel and also as a nonvoting member of the committee on physician programs, I think is what it was called at the time but that sort of oversaw some various programs including the diabetes provider recognition program which uses some of these diabetes measures.

Ann Hammersmith: OK, thank you. Ann Leddy? Is Ann Leddy on the line?

Katie Streeter: She won't be joining us today.

Ann Hammersmith: OK. Grace Lee? Is Grace Lee on the line? Laura Makaroff?

Laura Makaroff: Hi, this is Laura Makaroff. I am senior clinical adviser in the Bureau of Primary Healthcare at HRSA, the Health Resources Services Administration. I have nothing to disclose. There often is report on some of the measures including the diabetes control measure but I have not influence over those things.

Ann Hammersmith: OK, thanks.

Laura Makaroff: Thank you.

Ann Hammersmith: Anna McCollister-Slipp? Is Anna McCollister-Slipp on the line? Patricia McDermott?

Patricia McDermott: Yes, I work at Aetna and I'm director of clinical measures and operations here within the office of the chief medical officer. I use these clinical measures for provider performance but I have no other disclosures.

Ann Hammersmith: Thanks. Janice Miller?

Janice Miller: Yes, hi, I'm a nurse practitioner. I'm a certified diabetes educator and I'm an assistant professor at Thomas Jefferson University and I have nothing to disclose within the last year, within the last 12 months.

Ann Hammersmith: OK, thank you. Claudia Shwide-Slavin?

Claudia Shwide-Slavin: Hi, yes, I'm here. I am now working with Tandem Diabetes which is an insulin pump company as a clinical diabetes specialist in New York. And, I am no longer have the conflicts working with some of the other companies like Eli Lilly.

Ann Hammersmith: OK, thank you. Janet Sullivan?

Janet Sullivan: Yes, I'm the vice president and medical director of the Center for Regional Healthcare Innovation at the Westchester Medical Center. And, we are implementing on a countywide CMS project that will use these measures partly to evaluate a project and we will develop measures to evaluate our partners using some of these measures but I have not been involved in any of the development of endocrine measure.

Ann Hammersmith: OK, thank you. William Taylor?

William Taylor: Hi, it's William Taylor. I'm a primary care physician and a medical educator at multiple institutions in Boston and I have no conflicts to disclose.

James Rosenzweig: Yes, how do I mute that?

Ann Hammersmith: Thank you. Has anyone joined the call since we started and needs to disclose?

James Rosenzweig: Yes, this is James Rosenzweig. I was on the call but I couldn't figure out how to unmute myself. So, I was trying to answer you with you hearing me.

Ann Hammersmith: OK.

James Rosenzweig: I have no conflicts of interest to disclose but I am the chair of the Quality Improvement Subcommittee of the Endocrine Society and the Endocrine Society does create measures but none of them are relevant to this discussion today.

Ann Hammersmith: OK, thank you. Anyone else who needs to disclose?

Ingrid Duva: Yes, this is Ingrid Duva. I'm sorry. I was on the line and (I had) conflict between the computer problem. But, I have nothing to disclose. I'm still with the Atlanta VA Medical CenterStudied, still working as a fellow of the National Quality Scholars Program. So, we use the measures to help with implementation and to guide our quality movement. Thanks.

Ann Hammersmith: OK, thank you. Anyone else? OK. Before I leave you, I just want to remind you of one other very important thing which is that we rely on you as community members to work with us in the conflict of interest process.

And by that, I mean that if you think you have a conflict, if you think a fellow committee member has a conflict or if you're not sure or if you think someone is acting in a biased manner, please do speak up. We don't want you to sit down in silence. We want you to speak up in real-time.

You can do that by speaking up in a meeting or if you don't want to do that, you can contact your co-chairs who will contact interest staff or you can contact interest staff directly. And lastly, do you have anything that you want to discuss with each other or any questions of need that is on the disclosures this afternoon?

OK, thank you.

Katie Streeter: Great, thank you, Ann. So, this is Katie. I just want to remind everyone. We will be voting and discussing two measures today. This is a little bit different than our last webinar because we had a voting via a SurveyMonkey tool that occurred after the call.

So, we will be doing the live voting. As we move along, you'll see the slide pop up with the criteria that you're voting on. If anyone has questions, feel

free to send me a chat or we can – or just speak and pause and we are here to help you with that.

James Dudl: So, Karen, is the voting going to be verbal or is it going to be by the web?

Katie Streeter: We will be reading off what you were voting on but the votes will be submitted using the web tool. And actually, (Shawn), if you want to go into a little more detail about that now before we dive into the measures, that would be helpful.

(Shawn): Sure, Katie. The voting slides will appear directly in the slide window on your screen. You'll have small boxes next to each one of the options. You'll just simply click in the box of your choice so that we can register your vote. You do have the ability to change your vote. We just ask that you do it as quickly as possible so it gives the votes a chance to settle in so we can get the appropriate count.

If for some reason the box has failed to appear, next to your selections and you are a voting member, please refresh your session by pressing F5 or Command R if you're using a Mac. Non-voting members, you may see a box but it will not – you do not have voting privilege.

James Dudl: And just to be clear, for all the members of the committee, if they signed into the meeting using their invitation, they should be automatically able to vote.

(Shawn): That is correct and those folks that were unable to use the invitation or just didn't have it available to them at the time that they logged in, we have made adjustments and allowances for them in the backend.

James Dudl: OK. So, we have – so everybody should be lined up and ready to go.

(Shawn): Yes, sir.

Male: Great.

James Dudl: OK.

Male: Super.

William Golden: At that point, Jamie, my feeling is just sort of start doing the measures unless you have some opening comments you want to add?

James Dudl: No, I think that's fine. I think we should start with our first measure. Yes.

Katie Streeter: And actually, before we begin, I would like to give our developers a quick opportunity a few minutes just to give a brief introduction of the measure. We're starting with Measure 0061 and so if we have developers from NCQA on the line, please feel free to give a brief introduction.

Dan Roman: Hello. This is Dan Roman with NCQA. As a developer, this measure is the blood pressure control measure for our diabetes, our overall diabetes measure. The goal that – the blood pressure goal that we have here is less than 140 over 90. I believe that this is – and it is – it's looking at patients 18 to 75.

William Golden: OK, if that's the end of your presentation, I guess we should start with I guess Tracey Breen and Jim Dudl. We're looking over all of this.

James Dudl: OK, this is Jim. Can you hear me OK?

William Golden: Yes, I can hear you.

James Dudl: OK. And, I'm going to speak up for the moment but I've discussed this with Tracey and I'll start with comments but I would hope that Tracey will bring her view in also and if there are any questions, I hope you will ask them to either of us or both of us.

William Golden: Jim, one final comment. I talked to our staff and they basically advised us so we're going to do each measure individually and then after we have discussed these measures, if they get both approved, we would discuss if one should supersede the other because of incorporations.

So, we should take them as separate entities before we talk about comparable or how the inventory should work. So, I'm sorry to interrupt you, Jim.

James Dudl: No, no …

Tracey Breen: And, Jim, Tracey is here. So, take it away and I'll jump in if needed.

James Dudl: Please do, OK. Let's go on then to getting right on to it and that would be item number 1 and that would be the brief measure information. Obviously, this is comprehensive blood pressure control in people with diabetes and the target less than 140 over 90. It draws heavily, I would say from what's written and what we see from JNC 8, giving us so they actually need opportunity to review their process.

They of course started out with the condition of being evidence-based and – but also interested in practical issues of implementation. So, we will follow that. I don't know if you want to go right to the importance of the measure, the committee, pre-evaluation comments because I think we can cover the points as we go if you'd like to do that or I can cover them right from the start if you wish, whichever.

I will – actually, let's go with A1 evidence. For the moment, let me just cover that one separately. There was question for the committee, does the evidence indicates control blood pressure reduces undesired health outcomes? And, while the evidence was self-rated accurately as moderate because the exact goal of 140 over 90 wasn't – was not used. The quoted three RTC evidence – RTC's evidence clearly showed good quality and consistency of about a 50 percent drop in morbidity and mortality, controlling blood pressure below 150, although studies were somewhat older.

There are two other quoted studies that show basically no significant benefit by going below 140 over 90. And, the – as we go on then to our comments, I would say a couple of things. First of all, this is an intermediate outcome when we're looking at the strength of evidence. But, having said that, the other issue that they brought up is, you know, it's a very good evidence under 150 but they went to 140.

And, part of that reason as they said was to make ease of implementation and as you know, the general population target was 140 over 90 under 60 years old. And, they suggested that for ease of implementation, they make that

standard. And, we are directed to use grade criteria and the grade criteria includes the issues of ease implementation.

So, So I think that does fit within the overall concept. So, although I think you would say if it were 150 over 90, it would be high evidence given that this is an expert opinion, the overall evidence would be moderate. Going on to number 1B, the performance is gap, there was clear performance gap between 25th and 75th percentiles between 63 and 75, a 13, 14 percent gap. So that was quite consistent.

Karen Johnson: Hey, Jim. This is Karen. I'm sorry to interrupt you. Just because, you know, voting and stuff, let's just have a discussion about evidence first and then we'll vote on that and then go for it. Thank you.

James Dudl: Thank you. So, Tracey, anything you want to add to that?

Tracey Breen: No, I think that's a wonderful, you know, the summary I think we're dealing with intermediate and moderate issues that are, you know, people have kicked around these JNC numbers for a while. It would be interesting what other people in the group do into this. I think some of the comments are good but maybe people who didn't get a chance to submit comments can weigh in on thoughts now.

Sue Kirkman: This is Sue Kirkman. I didn't put this comment on this measure but it's just interesting that the other measure that we're going to discuss has the same blood pressure goal as a component of that measure and, you know, reviewing what I think is, you know, pretty much the same evidence, you know, they say that the evidence for blood pressure goal of less than 140 over 90, the quality is high and the strength is strong.

So, you know, I think it's – I just want to throw that out there as, you know, I think depending on how you define what the evidence is saying, you know, makes a big difference. So, my sense is that NCQA is looking specifically at the less than 140 over 90 or JNC 8, you know, looking at that exact target whereas my sense of the Minnesota people but develop the other measure is that they're, you know, sort of saying that the evidence for moderate blood pressure lowering is extremely strong and that, you know, we're saying that

less than 140 over 90 is moderate blood but not too stringent blood pressure lowering.

So, I just want to throw that out there.

William Golden: So, this is Bill Golden and I didn't submit written comments and I apologize, but I've submitted written comments on similar measures and I am going to vote to reject this on evidence and ask you all do the same for request of an edit.

In particular, there is no evidence that less than 140 over 90 is any different than measure than less than or equal to 140 over 90. But operationally, it's a huge difference. So, if doctors are seeing patients in the office and the patient has 140 over 90 and they record 140 over 90, they would fail this measure.

But if they wrote 139 over 89, they would pass the measure. There is no clinical difference in that but it does cause a tremendous amount of hassle as well as it would cause I think – it would undermine the integrity and the commitment to quality improvement measurements then people have this kind of (gumption) in the background on how they do quality measurement.

So, to me, this is a serious issue in the measurement community and that this measure really ought to read less than or equal to 140 over 90 and there is no evidence as far as I can see that would justify making it less than 140 over 90.

Sue Kirkman: So, I'm a little confused about – it seems like that's the case no matter what cut point you would choose. So, what is the difference between saying is that 140 or is that 139?

William Golden: Because in the office do rounding off in clinical operations and it is clinical – it is a clinical – if you're writing things down or typing things in, 140 over 90 is an easy round off which happens all the time in the office.

(Jesse): This is (Jesse). I was at – I'm impressed that you brought this up and wish I had done so. We actually did a study when I was at Hudson Health Plan and I don't remember the numbers but we ended up buying digital customer people because if they have a digital cough, they are likely to record, you know, if the

blood pressure is, you know, 139 over 89, that's what they'll put down, but if they're using a sphygmomanometer, they'll want to read it at that level of accuracy and people round to 140 over 90 even if it's just slightly below.

So, as we review, that 140 over 90 is often what's recorded and – yes, I agree, it's a picky detail that is a (gotcha) and not clinically significant. I think …

William Golden: Well, it could be a picky detail but it does – as I say, it will – it'll catch people and unnecessarily target are making tail measures.

(Jesse): Well, I agree and it's a picky detail that stands in the way of doing our thing.

Tracey Breen: You know, and just to throw out there and I don't want to hijack the conversation, this is Tracey, too much. There are real concerns about outpatient ambulatory blood pressure managements anyway. So, beyond rounding the validity of them and what they mean, you know, I think this is a murky area that it's something we measure but if we really want to get into the weeds on this, I think there's a growing concern about how valid outpatient blood pressure in the office measurements actually are.

So, I just wanted to throw that out there. Again, I think at some point, we have to draw the line in the sand about what we're going to measure and this is an exist – am I correct, Jim? This is an existing measure. This measure has already been in existence, correct?

James Dudl: Correct, yes.

Tracey Breen: Yes, so, you know …

James Dudl: It's existing but it needs to be changed.

Male: Bill, yes, I understand your point but by doing that, we're actually deviating some – I mean it's probably insignificant but we're deviating from the recommendations themselves, JNC 8 and I think ADA as well.

William Golden: Well, the recommendations are not sound for our measurement.

(William Taylor): This is (William Taylor), I'd like to agree with Bill Golden's plate and throw in another one, which is to at least raise the issue about JNC 8's conclusion that the blood pressure goal for people over the age of 60 should be a systolic of 150, not 140. We do have ADVANCE and ACCORD of the things that make us a little weary about hopefully aggressive blood pressure controlled for people with diabetes and it would seem to me, we might want to align with JNC 8 and instead of having a 140 over 90 or equal to or less than 140 over 90 for people over 60 to liberalize that to 150 over 90 to be consistent with JNC 8 and to prevent overly aggressive blood pressure reduction …

James Dudl: Yes, this is Jim. I think in JNC 8, that's in the general population 150 in the – people with diabetes, they did not make that statement. So, that would be a big deviation from what they say.

William Golden: Yes, but with – previous of recommendations for diabetics is like 130 over 80 and then that's been liberalized to 140.

Male: Yes, that's what kills people.

James Rosenzweig: Yes, that was in JNC 7 and JNC 8 because of the lack of good evidence in between 130 and 140, so that was why it was changed.

Male: Well, it's not the lack of good evidence. It's the presence of good evidence that the more aggressive blood pressure reduction ended up with the increase in mortality.

Sue Kirkman: Wait. What's there you're talking about because that is not what the ACCORD blood pressure study showed?

Male: To go back and do some ADVANCE or ACCORD.

Sue Kirkman: Well, that's glycemic control or the ACCORD blood pressure study tested a systolic less than 120 versus a systolic less than 140. And by the way, diastolic less than 80 in both groups and did not show any increase in mortality. It did show an increase in adverse events with the lower blood pressure goal but I don't know how accord blood pressure could possibly be used to justify saying it should be less than 150 as opposed to less than 140.

Male: Correct.

Male: I'm going to have to go back and do my homework, I'm sorry.

James Dudl: Are there other comments on this issue or do we go to voting I guess the last – (inaudible) of comments?

Katie Streeter: I think it should be less than …

James Dudl: I take silence as a time for voting. Is that correct in queue for the folks?

Katie Streeter: I think so, we can move on to voting. Kaitlynn is pulling up our first voting slide, I'm sorry?

Karen Johnson: This is Karen. We may want to see if NCQA has any comments about the suggestion to change it from less than 140 over 90 to less than or equal to 140 over 90 just to see what their thoughts on are.

Dan Roman: Sure, this is Dan. Thank you for giving us the opportunity. We did based our – this update of this measure which is a longstanding endorsed measure on the most recent JNC 8 guideline recommendations. As far as picking a threshold other than 140 over 90, I think somebody already said the recommendation is less than 140 over 90.

I completely hear what is being discussed as in it's kind of arbitrary but typically with measure development where we do have to draw a line somewhere since the guideline recommendations is less than 140 over 90. We would not arbitrarily choose to pick greater than or equal to – I'm sorry, less than or equal to 140 over 90. It's fairly common that we would follow the guideline recommendation when it specifically says a threshold which in this case for JNC 8 is less than 140 over 90 for the general diabetic population and that is kind of help – I mean that is how we can target our decision for what we put in this forum.

James Dudl: Thank you.

Katie Streeter: I guess, Kaitlynn, you're going to explain or maybe (Shawn) is if we need anymore additional information about voting but I will go ahead and tell you right now on your screen for evidence, you have five options. We're going to take the fifth option off the table. So, actually, what you have in terms of what you can vote for is high, moderate, low, or insufficient evidence.

If you feel like the evidence is insufficient for the measure and you are interested in thinking about doing the correction, please vote for option number 4, insufficient evidence and then we'll have a separate vote to see if you would want to do the exception. So, what we found on the past if people get confused, when we offered the five options, so we're only going to offer the four and if enough people are interested in – I think the evidence is insufficient then we could have a separate vote for applying the exception.

James Dudl: So, is the screen ready for voting? I have no boxes on my screen.

Female: Agree, I have no boxes on my screen.

Male: There we go.

James Dudl: Now, I do.

Katie Streeter: OK. Voting should be open for 0061. We're voting on evidence.

Male: We just click it, do we press enter afterwards or is it efficient just to click it?

Katie Streeter: Simply click in the box with your mouse. You'll actually see the box checked for you and it will register your vote.

Tracey Breen: I'm sorry, this is Tracey. I don't see any boxes. I don't know what I'm doing wrong.

Katie Streeter: OK. Tracey …

Tracey Breen: I see it, never mind, never might.

Katie Streeter: All righty.

OK, it looks like we have 18 votes.

Female: So, there were two votes for high, 11 votes for moderate, one vote for low, three votes for insufficient evidence.

Katie Streeter: OK, so the measure passed as evidence and we will move on to 1B, gap and care opportunity for improvement.

James Dudl: OK again, I'll take the lead and ask Tracey to fill in after. And this is performance gap and the opportunity for improvement as I started to say, there is a gap between 25th and 75th from about 63 to 75 percent, a 13 percent gap is a significant performance gap.

The second question for us was is there disparities data and as the application says that there is other evidence of disparity gap which I think is fairly strong but the submitter did not collect that kind of data. So, I think the significant thing would be that the evidence gap is clear and we will go on to the conclusion in a moment. Do we do priority at this point, too?

Katie Streeter: Actually, Jim, I meant to mention, that is another new thing here at NQF. We're no longer voting on that criteria so we don't actually have to have a discussion on that.

James Dudl: We will just skip it then.

Katie Streeter: Yes.

James Dudl: So, do we vote then on 2A separately or do we go on to reliability also?

Katie Streeter: I think we can pause and vote on 1B, the gaps in care or open up to the committee to see if anyone had anything else to add to the discussion.

James Dudl: And, Tracey, any comments?

Tracey Breen: No, I think this is a less controversial point than the last that we just made so …

Karen Johnson: Karen. Let me just make sure, even though we – the developer did not provide disparities data from their own measure on – the disparities question is really for just (archiving) purposes and also potentially another way to illustrate gap. So, they have illustrated gap with their performance right. So, don't let the fact that they didn't give their own measure information on disparities. Don't let that in particular influence your vote on gaps.

Katie Streeter: OK, it sounds like we can move on to voting. And voting should be open. Again, this Measure 0061. We're voting on 1B, performance gaps.

It looks like we have 16 votes. Up to 17 …

Male: How many people are voting?

Katie Streeter: We have 18.

William Golden: Perhaps somebody has dropped off the line. Why don't you give them another 10 or 15 seconds and I think we have a clear voting pattern here.

Katie Streeter: I think so.

Female: Katie, we do have 17 confirmed right now still online.

Katie Streeter: OK. So, for 1B, performance gap, 15 voted high, two voted moderate, zero voted low and zero voted insufficient.

OK, I think we can move on to scientific acceptability.

James Dudl: OK, then on 2A, reliability, the specifications are provided for administrative data and there was questions whether or not the blood pressure is clearly defined and I think everybody would agree with that. The update included the new diabetes medication, so that takes care of any other loose ends.

The logic is very clear and tested and it – is it likely to be consistently implemented based on prior use, I think yes. Do we go to 2A too or do we vote?

Katie Streeter: Keep going and talk about testing.

James Dudl: OK, then testing, we look at signal to noise and as you know one is perfect, a value of 0.7 is regarded as minimal acceptance. The signal to noise for health plan is 0.97 Medicaid, 0.95 for Medicare, very, very good. At the physician level, it's a medium reliability. It said on the sheet here 0.6 but the overall was 0.7. So, I would say the level is high for health plans and adequate for physician level.

Do we go on to validity or do we stop there?

Katie Streeter: Let's stop there and have a vote or open discussion for the committee and then go ahead and vote.

William Golden: So, I just want to be clear here, the reliability would be referring to the results in the record can be reliably extracted? Is that what reliability here refers to?

Karen Johnson: Bill, this is Karen. What they showed here is the score level reliability, so that is a measure of how capable is this measure of distinguishing between physicians. So, it's not in the data element and reliability that you just described. And, for these kinds of measures, we – excuse, we allow developers to show results for either data element testing or score level testing or both and they chose to do the score level testing.

William Golden: Then, my comment would be that it would be diminished by the rounding errors. So, probably is still useful but less so because of the inclusion of the less than as opposed to less than and equal.

William Curry: This is Bill Curry. One of the issues that I have when we look at these measures within our own practices, if we take the data from the electronic medical record and the blood pressure that is measured at the time, the patient is brought back to the exam room.

If that reading is elevated and it is rechecked during the visit either by a nursing team member or by the provider and it's found to be normal, there are many times when that improved blood pressure because the patient has had a chance to catch their breath is recorded in the provider's note but it's not

recorded as a common data element in the EMR many times for the data collection.

So, I think that provides an issue or problem with the data collection here. You know, we have tried to have our teams to recheck the blood pressure if it's elevated and if it's different than the first blood pressure reading to enter that second reading in and it'll be the most recent blood pressure then. But, I think it does impact on the reliability as this can be reported out of the EHR.

James Rosenzweig: Yes, I had a very similar situation. The initial blood pressures were done on, you know, on an automatic machine and a large percentage and were redone by the physicians but the actual value that went into the (M.R.) was the initial one.

William Golden: Yes, I have to agree. I – hopefully, I have not any technical problem. And I often see techs doing the intake and putting blood pressure cups over sweaters and things and I always will take my own blood pressures in the office but when I suggest that to trainees, I kind of feel like I'm a dinosaur.

Are people still on the line or have we lost contact?

Female: We're here.

Female: I'm still on.

Male: We're here.

(Crosstalk)

Tracey Breen: Again, it's Tracey weighing in the elephant in the room. It's very hard to discuss in a call like this is how accurate ambulatory blood pressures are. And again, because of the discreet data fields, we're using a lot of data mining. It's concerning but again, it's what we have and we're all using it and one would hope that the error, we're kind of continuing making, well, maybe wash out with the noise.

But, you know, it's a concern. It's also a continuum measure, right? So, you can check blood pressure all day long, which one is the true blood pressure. But, I think we're kind of stuck with that for right now.

James Dudl: Incomparability from group to group, physician-physician unless that changes quickly, it should be accurate.

William Golden: I didn't check with it. Maybe hear from the NCQA, if the (specter) is the last blood pressure taken as opposed to a trend line. So, if the patients had four visits and three are controlled and one is not, he would fail the measure if the blood pressure on the last visit is high.

Katie Street: Yes, what's the timeline?

Dan Ramon: So, this is Dan from NCQA. The measure is the most recent reading taken during the measurement year. And if there are multiple readings taken, physicians or clients can see the lowest readings that was taken and the lowest – sorry, yes, the lowest systolic and the lowest diastolic and reports that as the reading for the measure.

(Crosstalk)

James Judl: You mean the systolic from one visit and the diastolic from another …

Dan Ramon: No, no, sorry – this is – if there's more – there are multiple readings taken in one day or one visit, they can take the lowest systolic and the lowest diastolic and reports that. It's from the same day though.

Tracey Breen: So again, just be clear, over time, if someone has multiple touch points over a six-month time, it's the most recent visit that's looked at would be – I don't want to say the lowest blood pressure data points of that visit. Is that correct?

Dan Ramon: That is correct, yes.

William Curry: The challenge for those that collect the data however, this is Bill Curry. The challenge with those that collect the data when we do the data mining, we try to get the datasets to report these. It's the most recent that is – that's always pulled. I mean just it would be hard for – it could be done but it would be

hard to have an algorithm for our EHR to look at the last day that the patient was seen and they go through and pick out the lowest systolic and diastolic.

So, I think that, you know, it's going to be the most recent blood pressure that's going to get reported for these measures as we report them currently.

William Golden: OK, we hear no other comments. Shall we throw up the voting screen?

Katie Streeter: OK, voting should be open. Voting on measure 0061, reliabilities. And we have 17 votes.

Female: So, for 2A reliabilities, three voted high, 14 voted moderate, zero voted low, and zero voted insufficient.

Katie Streeter: OK, passed. Let's move on to validity.

James Dudl: And, I would say on validity, there were three different groups that provided face validity, one NCQA and two other expert groups. They all trained groups unanimously, said there was face validity. I don't think there's anything else in that component. Threats to validity, they did suggest excluding people with policies to go over a disease and just patient with steroid induce diabetes.

They did seem to run their data both with and without getting a little difference or no significant difference. So, it seems like there is no threat to validity by those exclusions and no groups were inappropriately excluded and I think that's about all we would have to say there. Meaningful difference of course, 13 percent less meaningful.

I think that's all I have to say. (Terry)? Any other comments?

William Golden: So, actually my comments from the first episode, I think the (face) validity, it does not adequately distinguish or fairly distinguish quality measures because of the less than as opposed to less than and equal. I will be voting as an insufficient and I hope others will join me.

Sue Kirkman: So, this is Sue Kirkman. There was more face validity presented on the validity testing. There were also comparisons to other quality measures,

showing a moderate to high correlations with multiple other quality measures but I believe there's significantly more than just face validity presented here.

William Golden: Yes, I'm sorry that's correct. And Measure 18 for example, yes, I didn't mean to short – change that. Thank you.

Katie Streeter: Are there comments from committee members? OK, let's go ahead and vote. Measure 0061, we're voting on validity.

(Off-Mike)

Katie Streeter: Looking like we have 16 votes. I know that we're up to 17, OK.

Female: For 2B validity, six voted high, eight voted moderate, zero voted low and three voted insufficient.

Katie Streeter: OK, the measure passed is validity and we can move on to feasibility.

James Dudl: Under criteria 4, usability and use, how usable is this and first of all, it is publicly reported and there are five groups including HEDIS, ACO Accreditation, NCQA Diabetes Recognition Program, using it – it looks like the use in the future will be greater as we use electronic medical records more that were no unintended consequences reported.

So, it looks like it's usable. (Terry)?

(Terry): Yes, I agree. Thank you.

Katie Streeter: We're actually – if we could back up to feasibility first and then we'll go ahead and – yes, I think we've skipped that one and then we'll go ahead to usability.

James Dudl: OK, yes, the measures collected through administrative data and so the questions for the committee, where are the elements routinely generated and used? It's proven collectible. It was ongoing audits by NCQA to be sure it's adequate and it's available on electronic form most of the places and likely that will increase, sorry.

So, I would say, yes it's – feasibility is high.

William Golden: I think you meant to say it's clinical data not administrative data.

James Dudl: I'm sorry, clinical data, thank you.

William Golden: So, as we do this, I just want to make clear, are approving the measure framework? Are we approving an eMeasure? Is that a separate consideration or is that basically eMeasure versus audited measures the same – and rolled up into the same discussion?

Karen Johnson: This is Karen. I want to ask the developer but my understanding is this was not submitted as an eMeasure, correct? So, this is …

James Dudl: That is correct.

Karen Johnson: OK.

James Dudl: That's correct. Yes, that's correct. It is not an e-measure.

William Golden: The only reason I asked that, I have heard tales of issues about harvesting vital sign data by HIE systems, so I just want to be clear there may be more technical issues there, one would expect for something that's basic so …

Katie Streeter: Any other comments on feasibility? OK, we're done. I think we can move to voting. Again, we're voting on 0061, feasibility. And, we have our 17 votes.

Female: So, for feasibility, 14 voted high, three voted moderate, zero voted low and zero voted insufficient.

Katie Streeter: OK, now, we can move on to usability and use. And Jim gave us a nice summary of that. Do we have other comments from committee members?

(Off-Mike)

Katie Streeter: And there's no comments. I think we can open up the voting.

And we have 15 votes.

Sue Kirkman: So, I just – this is Sue Kirkman, the developer does present data that there hasn't really not been improvement over time so – I mean I guess everybody has voted but I'm surprised that so many people voted high. Just a comment.

Katie Streeter: Karen and the co-chairs, are we acceptable with having 15 votes?

Karen Johnson: It looks like 17th one just came in.

Katie Streeter: OK …

Karen Johnson: So, I'm not sure if that's just people are a little slow on voting or if that's just our self that are being a little slow. But, I am curious. It is an interesting question that Sue brought up. So, I'll have you go ahead and read out our accounts and then maybe we could have just a quick maybe one or two-minute discussion about the improvement aspect of this. Just because it is one of our criteria and it's one that often doesn't get discussed very much.

So, I think it would be quite useful to have that discussion.

Female: For usability, 12 voted high, five voted moderate, zero voted low and zero voted insufficient.

Sue Kirkman: So, this is Sue Kirkman again. I mean, you know, the developers show that, you know, there really has not been improvement at least over those three years and that's consistent with CDC data for example that, you know, blood pressure control is one metric that doesn't seem to be really improving in people with diabetes whereas most of the other metrics are.

So, I mean that's – I only voted moderate for this but, you know, I mean the vote is in so …

James Dudl: But in the physician recognition program, they got up to 80 percent and in some groups, there's 80 to 90 percent. I don't think it's the measure that's the problem, OK? You know, it's what people do with the measure and certain people seem to be able to make it work.

Sue Kirkman: I didn't see that diabetes recognition program data. Was that presented here?

James Dudl: I think it was and I think it was 80, 82 percent. NCQA can comment.

Sue Kirkman: OK, under usability and use, all I see is the HEDIS plan data.

William Golden: For the sake of expedition, we can take some of this offline as a moderate vote and a high vote would basically still pass the measure. So that's …

Sue Kirkman: I think.

Karen Johnson: And this is Karen. Yes, the 80 percent numbers actually came up under 1B under opportunity for improvement. That's where they put that data.

William Golden: I think it's now time to go on to the next item, endorsement, is that right?

Katie Streeter: That's correct. Any other discussions from committee members or are we comfortable voting?

William Golden: I know I've been in the minority but I cannot endorse it.

(Off-Mike)

Katie Streeter: And voting is open. Overall suitability for endorsement of measure 0061.

(Off-Mike)

Female: And we have 17 votes. It looks like 14 voted yes and three voted no. And that will also include our discussion for 0061. I think we can move on to our next measure.

James Dudl: So, it's either Bill Curry or Sue Kirkman and up next.

William Curry: I'll start and Sue will add on it if that's OK.

Katie Streeter: I'm sorry, just to jump back in, I'd like to give the opportunity to introduce the measure before we have our discussion.

Collette Pitzen: Great. Thank you. Good afternoon. I'm Collette Pitzen, a measure developer with Minnesota Community Measurement. And with me is Jasmine Larson,

manager of measure development. We're pleased to present – be presenting Measure 0729, optimal diabetes care for consideration of re-endorsement.

This measure of the patient level all are non-composite measure that seeks to reduce modifiable risk factors associated with long-term macrovascular and microvascular complications associated with diabetes. Diabetic patients are more likely to reduce their overall risk, prevent or reduce complications and maximize health outcomes by achieving several intermediate physiological targets.

This measure is specified for recording at the medical group and clinic site level. The denominator is adult patients age 18 to 75 who have a diagnosis of diabetes and are considered established to the practice by having at least two face-to-face visits with diabetes ICD-9 codes in the last two years and at least one visit for any reason during the measurement period.

Exclusions are as follows. Permanent nursing home residence, hospice or palliative care, death, pregnancy during the measurement period and diagnosis coded in error which is frequently pre-diabetes. The numerator is calculated at the patient level and numerator compliance is defined as the patient achieving all five components of the measure which are hemoglobin A1c less than eight, blood pressure less than 140 systolic and less than 90 diastolic.

Statin use is no contraindications or exceptions, tobacco-free and daily aspirin or antiplatelets if the patient has ischemic vascular disease and no contraindications or exceptions. The components are treated equally. There is no waiting of the individual components.

This measure originally developed by HealthPartners have been publicly recorded in Minnesota for over seven years and have been included in the Accountable Care Organization measures at the least three years.

The components have undergone slight changes over the years related to changes in evidence and guidelines but have remained fairly stable until the recent paradigm shift in cholesterol management guidelines. The cholesterol component was previously defined as the target of LDL less than 100.

However, based on current guidelines and evidence, this is no longer an appropriate component or target.

In 2014, the cholesterol component was redesigned to reflect appropriate statin use for this population. This change in the measure will be effective for 2015 days of service recorded in 2016. The measures included in the Minnesota Department of Health Statewide Quality and Recording Measurement System which requires patient level data submission for all primary care and endocrinology clinics in Minnesota.

As our group's CRM systems mature, more groups are able to submit their full populations and currently on an annual basis, we're receiving data on 97 percent of the eligible population within the 2014 reporting year, represented 118 medical groups, 580 clinic sites and over 230,000 patients.

So, the 2014 reporting year, the statewide average is 38.9 percent of patients achieving all site targets, increased from our initiate rate in 2000 or 2006 of 9.5 percent. Although the measure has a period to plot to over the last four years, the number of numerator cases have increased by 30,000 more patients meeting all targets. And there are significant variability among the practices with the lowest scoring clinics at 5 percent and the highest scoring clinics at 78 percent.

Interestingly, we also see quite a bit of variation between clinics within a medical group. This measure still demonstrates the gap and opportunity to improve. Please note that while the all or none composite is considered to be the gold standard, the individual components may be measured as low and are particularly helpful in quality improvement efforts to better understand where opportunities for improvement exist.

The statewide rates for the individual components are 73.7 for A1c less than 8. 63.7 for LDL less than 100. 84.3 percent for blood pressure less than 140 over 90. 99.5 percent of daily aspirins if a diagnosis of ischemic vascular disease is present and 84.5 percent for tobacco and nonuser.

We thank you for the opportunity to present this measure for your consideration and I additionally like to thank the committee members for all

the great and thoughtful comments that were shared with us prior to the meeting and we are happy to address any and all questions that you may have. Thanks.

William Golden: I have a question for the developer. I believe the Diabetes Association has recommended statins for patients, diabetics over the age of 40. Your measure would have at 18 to 75. You want to comment on that difference?

Collette Pitzen: Great. Thanks, this is Collette again. A great question. We have actually built and this is probably buried in the materials that you received. We built the ACC/AHA guidelines in to the calculations for the statin component because sometimes, a diabetic patient at a younger age may actually be at high risk for example if they have an LDL greater than 190, that kind of changes.

So, the workgroup actually had a lot of thoughtful considerations in addition to patients having diabetic or having diabetes and ischemic vascular disease. If a patient has ischemic vascular disease, the age range is 21 to 75. So, we took all of that into account into the algorithm for calculating the statin use.

James Rosenzweig: This is Jamie Rosenzweig. Did you take into account appropriate statin use like low dose, medium dose and high dose according to with the recommendations?

Collette Pitzen: Sure, I'd be happy to talk through our workgroup's process. We actually started out with a consideration for the diabetic patient being on at least the moderate statin. We started our development work shortly after the ACC/AHA guidelines came out. We also rely heavily on the institute for clinical system improvements and their guidelines.

And, we took a short pause for those guidelines which were published in July and then we completed our work again. And, in terms of burden and in accordance with suggested measurement. The workgroup decided that do not designate the level of statin of the patient (as on) simply are on the statin as appropriate for the algorithm.

And some thought in dealing from our cardiologist from a group that even on low level statin is going to have some benefit. So, at this time we are not requiring the dose in the consideration.

William Curry: This is Bill Curry. In the packet of information that the measurements (Stuart) presented, there is a PDF that shows a very nice flow diagram of how to determine if the patient meets the statin criteria and I think it's very well done. I also appreciate the fact that at this point they don't talk about whether it's a moderate or high strength statin but I think they've done a very nice job to make this very clear and it's a separate PDF that you can find on the NQF website for this measurement.

Sue Kirkman: Yes. This is Sue Kirkman. I agree and when I initially saw this, I thought it was sort of saying that anybody with diabetes over the age of 18 should be on the statin but if you look carefully at the specifications, you know, for people under age 40, you know, it's very clearly only if it's, you know, secondary prevention like they are – they already have a ischemic vascular disease or very high LDL. So, I think it's very appropriate and consistent with the guidelines.

Male: Are we ready to start with our discussion. And I think so Billy you up.

William Curry: Sure. We'll talk briefly about evidence to support the measure. So there is five measurements in this composite. The first is glycemic control and I there is really a high quality evidence and strong evidence to support that having glycemic control and their level of A1c less than eight would improve outcomes for a group of patients with diabetes.

The second is the hypertension control and it's a very similar to the conversation that we just had with the previous measure and it's the same cut points of one – less than 140 and less than 90.

The next piece of the composite is statin use and they have I think prepared a very nice diagram of how to qualify whether patients have meet the measure or not and I think that there is good evidence that and it's both the strength and quality that would support the treatment of patients with statin therapy.

The next is smoking cessation and it's tobacco free so it's a measure of whether the patient is smoking and they argue that there is plenty of evidence that would state that if we could help our patients to become tobacco free would be significant reduction and morbidity and mortality in this population.

And then the final measure is that of aspirin or antiplatelet therapy and they provide evidence that there would be benefit in this group to reduce complications with the use of aspirin or other antiplatelet therapy and I think there is high quality and strong evidence to support that.

One of the pieces that they did mention several places in their documentation is that if we as providers, clinicians spend effort to look at one or two of these efforts, I think that we can make some impact in our patient's mortality and morbidity into the future, but certainly if we're able to look at the patient as a whole and look at each one of these measures in the composite and positively impact them all, it will be even more powerful to reduce mortality and morbidity and they provide some evidence to talk to that. So, Sue?

Sue Kirkman: Yes. I don't have much other than the, you know, my comments about the blood pressure goal evidence, you know, being described as high and strong and this one versus expert opinion and the other one that probably interesting but …

Female: Can we go over?

Male: Again, the discussion here would be evidence in sense that would apply to an individual patient or a management of the disease correct not in terms of how it discriminates across practices. That would be a different category, correct?

Sue Kirkman: Right, that would be more reliability I think.

Male: Yes, OK.

Female: Yes, it's correct.

James Rosenzweig: Yes, this is James Rosenzweig, you know the endless discussions about where the cut off should be between, you know, for a population of patients

with diabetes between seven and nine percent but for the purposes of this composite measure it seems to me that eight percent reasonable one.

William Taylor: This is William Taylor. Can we discuss what is the right measure of quality is, where the patients tobacco free or whether the clinician is working with the smoking patient to help him or her quick?

Male: I was going to talk about that down the road perhaps in validity because I believe that's a phase validity issue with this the step part of the composite. And we can talk about it now. We can talk about it later.

So, my thought is that with this – with all of these measures (minors) the tobacco it's something that the provider has to initiate. They have to initiative a prescription for, you know, medication as to treat blood pressure. They've got to initiate medications for the lipid management but there is no ability to account for the multiple interventions that we might try with the patient to assist them in smoking association.

So, that is perhaps not very dependent upon the provider and much, much more dependent upon the patient more so than I think the other four measures in this composite.

Sue Kirkman: But in the way it's kind of the same thing is, you know, you can see a prescribing (site) blood pressure medication and the patient has just really difficult to control blood pressure whereas another patient, you know, may not have difficult to control blood pressure and, you know, so I think that's kind of always the case that there can be a lot of effort in the background that is not captured by these measures.

I also have kind of philosophical concern about all are none, you know, as a measure of a quality because it sort of implies that, you know, if I had all four – if I had four but not the fifth one, that's the same as if I'm not meeting any of the five, you know, and my patient and, you know, I sort of gets into, you know, more of a philosophical question but, if you have to meet all five and you have to cross, you know, thresholds. Some of them are continuous variables and you have to cross thresholds for several of them and you know,

it just seems like there can be a lot of gradations and quality in the, you know, people they don't pass the measure.

But maybe we're something ahead too much.

Karen Johnson: So this is Karen and under criterion 1D, how you will get to talk about this component of this composite so I think the discussion about (all or one) and are these the right things to pick in the composite. You'll have that for 1D but I just (hear) back to thinking a little bit more just about evidence for tobacco associations at this point.

Female: Very strong evidence that being tobacco free is, you know, associated with improved outcome. So, I don't the evidence for the recommendation or the, you know, the threshold is lacking.

Male: Is this composite measure listed as a measure of quality of care?

Female: Yes, we will consider this a quality of care measure.

Male: Some components are clearly more within the capability of the care giver to provide another (inaudible) composite measure.

Male: So may be we can vote on the evidence and then move on and talk about some of these other issues as they come up and later maybe parts of our discussion.

Female: Yes, that that would – and this is the four with the evidence criterion. You know, you have the wrong slide up, Kaitlynn, there you go. You'll be voting high, moderate, lower, insufficient, again, option number five will take off the table and if we need to go there we can.

William Taylor: So again, this is William Taylor help me again, so if we believe that having the patient to be a tobacco free is not proper measure of the quality of the care the patient receives since it's outside of the control of the clinician to do that, if the clinicians doing everything to help the patient become tobacco free and how do we – does that enter in to our view of the evidence?

Female: It won't come under evidence. That's more of a question about validity. So, is it a valid measure of quality? So you can discuss that under validity and vote

there, so this one is just simply on the clinical evidence underlying the component and we don't ask you to vote separately in each of the components. You know, we have to, the, you know, you think about all of them combined, you know, based on the evidence that's been presented.

Is that makes sense?

William Taylor: Yes.

Female: OK and we will open voting. Again, this is for evidence on measure 0729.

(Off-Mike)

Female: And that's 17 votes. So for 1A evidence, five high, 11 voted moderate, zero voted low, one voted insufficient evidence. And now we can move on to 1B opportunity for improvement.

Male: So the measurement (Stuart) had mentioned that on the individual component there was some fairly high compliance with the individual measures in the composite ranging from 64 percent from – for LDL control to a nearly 100 percent for – in the platelet therapy, yet over the last three years the composite score has plateaued just under 40 percent. So, it would show that if we were to put our patients – the patients from Minnesota in this group and ask how many of them met all five measures, it would be just a little over a third of them.

So, that shows that perhaps there's a gap in performance with helping our – this patient cohort to become best controlled. They also did present some evidence that there might be disparity issue here and that African-American residents in Minnesota did score less than the non-African-American population. They we're only able to get to that 27 percent of those individuals to meet the composite measure.

William Curry: Now, so that – I mean that discussion is interesting because it – perhaps it comes up into validity. But, you know, there's all this talk about socioeconomic risk adjustment. So if you are caring for, you know, population with increased burden of behavioral health issues, which have

higher smoking rate or different socioeconomic status where you have different rates of smoking. Now that kind of biases the outcome of the measure.

Not only where that would fit in or how we would adjust for that. I'm not sure under this category, but I think it does have a problem with how you interpret the data across providers.

Karen Johnson: And Bill, this is Karen, we'll talk about that issue specifically in two places, in reliability because we need to talk a little more bit more in detail about the specs and then – and validity. I think one of the things that I was a little confused about and we'll ask the developer to clarify for us. They talked a lot about risk adjustments (today). The only thing that I was confused about is are they actually doing the risk adjustment now or are they proposing risk adjustment?

But again, we'll push that off until the reliability and validity discussions.

Male: Yes, I would agree. I don't think it's under this topic.

Katie Streeter: And if there are no other comments, we'll go ahead and open the voting.

OK, voting is open for performance gap in measure 0729.

(Off-Mike)

Female: So whoever is talking to themselves you might want to mute your phone.

Male: Have some people left us?

Claudia Shwide-Slavin: This is Claudia, I sent in a notice that I had to travel starting at 3:30 so I can only be listening in, I can't vote then.

Female: OK, looks like we have 15 votes for high. And I think that would be enough to move on to 1C. Or actually we're no longer doing 1C. So, we can move on to 1D which is specific for composites, quality, concept and rationale.

Male: So, this is the five-component composite measure. It's scored all or none. So, the components are all modified but risk factors that are important to reduce the risk of morbidity and mortality of diabetes. And with being achieved together will significantly impact the patient's health. They're measured at the patient level and attributed to the provider who provides the care for the patient. And they don't wait the components.

I don't know if Sue wants to add any comments or ideas here.

Female: Yes. I thought it was reported at the health group or the group level as opposed to the provider.

Male: Well, I should say, it's measured at the patient level, it's attributed to the provider, and – but it's reported at the health group level, that's correct.

Female: OK.

M. Sue Kirkman: I would also – This is Sue. I have kind of a question for the developer that, you know, you mentioned that it's more patient-centric to look at a composite measure of multiple indices of quality versus looking at them separately. But I guess I'm a little unclear on why – I mean, these are five – you know, each of these components is performance measure in various systems. So, you know, there is a hemoglobin A1C left in a performance measure and there's a, you know, there's blood pressure measure and there are certain measures and so forth.

So, I guess I'm just a little unclear on why it's more patient-centric to sort of have an all or none, you know, you fail if you have only four of the five or if you have zero of the five versus looking at all of these things but in separate measures.

Collette Pitzen: This is Collette. I'd be happy to respond to that about our philosophy and belief.

Yes, you indeed could look at everything separately and we do have information about the different components. The thought is if one is only focusing and paying attention to transparency of measurement around A1C

and perhaps not taking into those other factors, you could look like you're doing it really well for managing the patient's blood glucose and their cholesterol management or their blood pressure could be sky high.

So, the philosophy is that by managing several of these modifiable risk factors together, it represents the best opportunities for an individual patient to prevent or prolong or stall long-term complications.

M. Sue Kirkman: Yes. I mean, I understand multiple risk factor modification, but I guess I'm not clear on why having five different performance measures versus having one, you know, composite all or nothing measure that incorporates the five different performance measures is more patient-centric.

(Steph): Sure, this is (Steph).

(Jesse): This is (Jesse). Maybe …

(Steph): … I could just add, you know, with five separate measures reported independently, a hypothetical practice could perform let's say 60 percent on all of those measures and potentially be looking like they're doing a fairly adequate job of managing their patients. However, in reality, they may not be doing a really good job of managing multiple risk factors with individual patients and the overlap of multiple modifiable risk factors for patients is not apparent.

(Jesse): I think that's also – This is (Jesse). Intuitively, to patients, how many of your patients are doing well? And it also – Doing all or nothing at the level of a patient also is one way of doing measures in practices where you have small numbers. So, whereas, you can't really say, you know, if you only have three patients, you can't really say that, you know, three patients with diabetes, you can't really say that your A1C rate is anything because you don't have enough patients.

But if you're doing how many of your patients have all of the measures that they should have, and, in this case, we're looking at diabetes measures, you can say that across a whole population of your patients. If you only have two patients with diabetes and you measure both of them, then you're describing

the entire population that you have. So you're not saying you're – Anyway, I think I've said it.

William Taylor: This is William Taylor. It is a measure that's patient-centric in the sense that it takes all the measures for one patient but it's certainly not the only way to do that. In all the controversy in the latest lipid recommendations about the risk calculator reminds us that there's been a lot of attention and how do you bring various risk factors together and try to make one prediction for one patient. You don't have to say you have to achieve all signs in order to have high quality care. What if you had high quality on four of them and you did pretty well on the fifth one, that's not the same as somebody who doesn't achieve any of them for all five.

So, did you say all five achieved or that fails is not the only way to capture looking at all five criteria for one patient. It seems to me there's also a danger if things are not doing – as I mentioned before about adjusting by socio-demographics. But if that's not done properly, then since a lot of these streets (frack) by socio-demographic characteristics, it could be an unintended consequence to me, practices less hospitable to the patients who are less likely to achieve all five which, you know, the people who achieve all five might track more with the socio-demographic status and those with the quality of care rendered.

(Jesse): This is (Jesse). I actually agree with that and I thought that the discussion of risk adjustment with this measure was very thoughtful. That way I – in (Karen's) comments about, I want to hear if they're actually doing it or just talking about it, but I thought that the way they approached that was – just hit the sweet spot. You don't want to adjust the way and say it's OK for poor people to have worse care. But on the other hand, you want to be able to say, "Those who are taking care of poor people have a harder job to do," and the way that this developer described reporting that you can say, "OK, your scores aren't that high but you're doing better than expected. And so I thought it was a very good discussion of how to do a risk adjustment in this case.

William Taylor: So, that was – This is William Taylor again. That was asked before. Are we voting on the properly risk adjusted arrangement or one where that's a hope that might someday be implemented?

Male: Well, yes, let me ask that question on what (Jesse) just said to maybe NQF staff or how this would look out. I'm sorry, I had to drop off. I had a computer glitch. I'm coming back here maybe.

When NQF endorses a measure, it would seem to be that the measure developer here is saying that there is an issue depending on the population being measured, but that the measure would be useful for that entity conducting the measure for internal benchmarking but it may not be appropriate for accountability across all providers.

Now, in the back of my head, that's not usually how NQF uses or has its measures being used which tend to be used to compare across providers. Now, how does that play out? Do we actually then put in some sort of a caveat to future uses of the measure that the measure is not intended to be used across provider sets?

Karen Johnson: So, Bill, this is Karen. Currently right now, NQF endorsement implies that measures that are endorsed are suitable for endorsement for accountability purposes, you know, be it public reporting or payment or other kinds of uses as well as internal Q.I. So that's one of the things when we get to reliability thinking about the specimen and particularly the validity, you ought to think about the risk adjustment approach and decide if you think that this measure is appropriate to – for both of those types of uses.

Now, in the future, NQF is going to be seriously considering changing what we mean by endorsement and think about endorsement – I think the term that we're using now is for intended use. So, perhaps we want to see different things for measures that are going to be using accountability programs versus internal Q.I., but we're not there yet.

So, right now, we are operating under our current policy which is measures should be usable for both purposes, internal Q.I. as well as accountability.

Female: And this is community measurement. If I may just add, I'm sorry if someone – if we misspoke her or if there's a misunderstanding that we do believe that the composite measure is appropriate for accountability and not totally quality improvement. So, just to clarify that.

Janet Sullivan: This is Janet. I think that this – this is just such a comprehensive measure to really give us an idea not on a patient level or provider level but on a health plan level and really looking at this as national data. I think that although we know that we've had similar data reported in the past for three measures like – I think it was A1C, lipids and hypertension, and they were appallingly low numbers or percentages of patients that were well controlled on all three measures.

So, I think when we see those appallingly low numbers, I think it really is a motivator to providers to make some improvements in our processes as well as in what we measure. And I just think a comprehensive measure like this is just really important and we haven't had large data about that for several years. So, I just think this is an important comprehensive look at it.

William Taylor: This is William Taylor again. Could somebody respond to the sort of potential unintended consequences of health plans thinking of doing the right thing and, you know, marketing to people who might need their services the most and, you know, disadvantaged people, and so on, and say, "Wait a minute, that's going to, you know, destroy our quality measurement on this one as we go forward," or, you know, individual providers who want to make their practice most hospitable to the patients who are likely to give them the best score, and so on. Is there any concern along those lines?

(Jesse): This is (Jesse). I just – I don't think that's a trivial question. I think as it – as the measure is being used in Minnesota, it's valuable. And because it's so widely reported, I don't think that there is that issue. But I don't think your question is trivial.

Karen Johnson: So this is Karen and I'm looking at our time and I'm realizing that we're not – won't be able to get through this entire process for this measure. So we are going to have to come back to our second call that we have scheduled for next

week. So, again, I don't want lose track of that question. I think it's something that we need to think about a little bit more.

So, I think, for now, what we need to do is go ahead and vote on 1D which is the quality construct and rationale. And basically, I think what you're voting on for 1D is whether or not you believe that the developer gave a logical rationale for why they did what they did. We will be looking at criterion 2D a little bit later on under validity where you'll actually look at some empirical numbers and think about how it's actually working. So, we've split it up into kind of the idea about it versus the results of it.

So, for right now, I'd like you to go ahead and vote on 1D and then we will come back next week and Katie will give you details about that, but we'll come back next week and finish this discussion of this measure.

Katie Streeter: And voting is open.

(Off-Mike)

Female: Could people mute their phones? Really, I mean, I think if you're talking to yourself or to other people about how to vote, it's probably better not shared.

Katie Streeter: And we have 15. I believe that's all we have voting at this time.

(Off-Mike)

Operator: That is correct, we have 15.

Katie Streeter: For 1D, we have four voted high, seven voted moderate, four voted low and one voted insufficient.

Female: Right. And just to (prime us) in thinking about things because I think our next call getting into the reliability and validity of the measure and talking about risk adjustment, et cetera, it's a little bit of a deep dive that we'll have to think about.

Developers, just real quickly, can you swash my curiosity and tell me, are you actually using risk adjustment now and that's how you're bringing forward this

measure or are you saying that at some point you want to go forward with risk adjustment? That was – There were a couple of statements that made me miss – I couldn't understand what you are proposing.

Male: Sure.

Collette Pitzen: Great. Thanks. This is Collette. We are risk adjusting this measure right now. We are changing methods, and that's what we tried to describe. So, the risk adjustment variables which do include insurance product which is a proxy for socioeconomic status in addition to age band and type of diabetes are some of the variables that are used in the model.

The measure is being risk adjusted. Our board has had lengthy discussion. Previously, they felt that unadjusted rates for our consumer-facing website were best, and that line of thinking is changing. So we've risk adjusted in several places including our hard copy quality report, but we are now moving to a new methodology and I want to say in the next month or so, we'll be live with this measure included that is going to an observed to expected model for doing risk adjustment.

Is that enough? But we have been risk adjusting this measure for several years and we are changing methodologies and how we will be displaying that to consumers.

Female: OK. So let me just rephrase real quickly. You are saying that what you put forward in the specs now is what you're going – you're moving to, and that's – so what's in front of the committee is what you actually are going to be doing a month from now.

Collette Pitzen: Correct. Correct.

Female: OK.

Collette Pitzen: We have been risk adjusting based on a case mix method prior to our …

Female: OK.

Collette Pitzen: … forward this spring.

Female: OK.

Collette Pitzen: Yes?

William Taylor: And William Taylor again. And for clarification, when you risk adjust, the only component that picks up socio-demographic factors is the insurance product, is that right?

Collette Pitzen: That is correct.

(Jesse): This is (Jesse). If I could just ask, I know we'll be discussing this next time. It was something that the staff brought up in their comments which is that, in your description, you described a number of variables that look like they might be appropriate for risk adjusting, but then when you – in your risk adjustment model, you weren't using all of those variables. So, when you present it to us next week, could you address that, how it is that you ended up with a subset of the variables that you were going to use?

Collette Pitzen: This is Collette. If you don't mind, I could quickly explain that. All of those variables were tested in the model and they didn't have high enough (tee) scores to go forward.

(Jesse): OK.

Collette Pitzen: Yes.

Katie Streeter: And this is Katie. I know we need to pause the discussion. We'll finish next week. But I would like to open up the lines for our public comments. If there are any people on the phone that would like to comment, please let us know.

Operator: Thank you. At this time, if you have a public comment, please press star one.

And there are no public comments at this time.

Katie Streeter: OK. So we have our second webinar scheduled for next Wednesday, the 28th from 3:00 to 5:00 p.m. Eastern Time. We'll be using the same webinar links

and the same voting links. However, I will send out a reminder e-mail so you have that information and we'll also make sure it's in the calendar invite.

And any other comments from our co-chairs or committee members before we wrap up today's call?

Male: No.

Katie Streeter: OK. Thank you all for your time and we look forward to continuing the discussion next Wednesday.

Operator: Ladies and gentlemen, this does conclude today's conference call. You may now disconnect.

END