NATIONAL QUALITY FORUM

Moderator: Sheila Crawford
June 5, 2012
2:00 p.m. ET

Operator: Welcome to the conference. Please note today's call is being recorded. Please standby.

Suzanne Theberge: Good afternoon, everybody, and welcome to the Neurology Workgroup Two Conference Call. This is Suzanne Theberge at NQF and I am here with my colleagues Karen, Reva and Jessica. And we have the Steering Committee numbers for Workgroup Two as well as several of the developers on the line.

I'd like to start just by asking the steering committee members just to introduce yourself so we can do a roll call. Just say your name and where you work and let's do it in order of the agenda. So we'll start with David Knowlton on 1955.

David Knowlton: David Knowlton, I'm a president and the CEO of the New Jersey Health Care Quality Institute and I'm a co-chair of the committee.

Suzanne Theberge: Thank you.

Ramon Bautista: I'm Ramon Bautista from the University of Florida in Jackson. I'm a neurologist.

(Carrie Richmond): Hi. This is (Carrie Richmond) from University of Pennsylvania and I'm a nurse.

Risha Gidwani: And this is Risha Gidwani. I am a health services research at Stanford University Medical Center.

Suzanne Theberge: Great. Thanks, everybody. Well, I think we can get started. I'm going to turn the call over to Karen Johnson, the Senior Director on the project to introduce everything.

Karen Johnson: Hi, everybody. This is Karen Johnson. Thank you for taking time out to talk to us today on the today. And thank you for all the work that I know that you have done in looking at these measures. You may have realized that as you will working through several of these measures, these are probably our most complex measures in the project because today we are talking about free outcome measures that brings in the discussion of risk adjustments and that sort of thing. So, these are pretty complex so again, thank you for your work and evaluating those measures.

So let me give you just a little bit of an idea about how we're going to go about today's call. We have roughly about 20 minutes per measure that we can talk so those 20 minutes will go really fast. We've asked one of you to be what we're calling the lead discussant for each measure. So, as we discuss and what we'll ask you to do is just very briefly introduce your

measure and then once you've done that, go ahead and just roll out into summarizing your thoughts about the initial evaluation result around importance.

And once you do that then we'll kind of open up the call for the workgroup members to discuss points about importance and that would be impact, gap, or evidence. And once we kind of get through that piece of it then I'll ask the lead discussant to go back and talk about the results from the preliminary evaluations from the (discussability) portion and so on.

So that's certainly the idea and we'll just work our way through all the measures. As you know the developers are on this call. So as things come up, the things were unclear or you had questions or whatever, developers will be able to hopefully answer any questions that you might have. And our role here for – two of our main roles is really just to clarify if you need any thoughts and our criteria or our guidance in making (inaudible). So, we will be doing that. We'll also be keeping an eye on the time.

So, if the – we'll just your forgiveness if we have to cut you off in a really nice discussion but we may need to do that just so that we can be able to cover all the measures on today's call.

And then finally, Reva Winkler is here. We're very lucky to have her here, she's going to help us facilitate this call. So, thanks, Reva, for coming through our call.

So with that, let's go ahead and get started. Our first measure is measure …

Reva Winkler: Sorry, Karen, do you mind if we get the steering committee members get a chance to hear (what) the developers are actually on this call before we go ahead?

Karen Johnson: Oh, sure.

Reva Winkler: Thanks.

Karen Johnson: Developers, you want to introduce yourself?

Susannah Bernheim: Susannah Bernheim with the Yale team is here for the (today's) measures.

(David Seymour): (David Seymour) from PCPI is here for the CT MRI measure.

Deidra Joseph: Deidra Joseph and other AMA-PCPI staff is here.

(John Batt): (John Batt) with AHRQ in relation to the AHRQ mortality rate measure.

Female: (Inaudible) with American Stroke Association with Dr. Schwamm, (Dr. Faneroy) and (Dr. Smith) I believe as well as maybe in Nadine Allen is also ASA staff.

Judy Burleson: Judy Burleson, American College of Radiology staff.

Karen Johnson: Great. Thanks, guys, very much. OK. So, our first measure is measure 1955, the NIH Stroke Scale Reported. It's an AMA-PCPI measure and Dr. Knowlton, this is your measure, so would you like to introduce the measure for us?

David Knowlton: Sure. The NIH Stroke Scale is a greatest neurological examination consist of speech, language, recognition, inattention, visual feel abnormalities, motor and sensory impairments and ataxia, takes less than five minutes to perform. It's a process measure that if adopted would apply to the percentage would be interpreted as follows.

The percent of patients over 18 with ischemic or unsatisfied stroke with initial NIH stroke scale recorded. It's endorsed by the American Heart Association, the American Stroke Association. (Inaudible) folks from the NQF staff – do you want me to go into conclusions from this, or how do you want to – I'm sensitive to your time.

Karen Johnson: Maybe just ahead and talk about the first criteria and which is important to the measure and report, so just that part.

David Knowlton: Well, it's – as I mentioned, it's endorsed by AHA and American Stroke Association about 795,000 people experience stroke each year with an estimate $34.3 billion cost in 2008 down $8. Previous studies on this scale emphasize disability to predict mortality in acute ischemic stroke. Although there were some limited sample size in some of this study.

A 2012 study that was done to look at the relationship stroke scale in 30-day mortality and Medicare beneficiaries concluded the index strokes regarded a very strong discriminator of mortality risk even on the absence of other clinical information with the (inaudible) continuous of categorical risk determinant.

The patients have – about 2/3 of patients have documented stroke scale reported. It's a use of guide to treating strokes mainly with the tPA decision. The type of evidences is a clinical practice guideline. And they have done a number of studies that have demonstrated its ability and liability.

Overall, the agreement between nurse and neurologist relative to the items was not significantly different from agreement between neurologists. Trained research nurses can administer to the reliability similar to stroke-trained neurologist from a study that was done in 1997.

The grading of the strength of the quality of body of evidence, AHA and the ASA truth guidelines created the evidence plus one condition for which there is evidence for and our general agreement that have given procedural treatment is beneficially useful and effective.

Its level of evidence which means the data was derive from a single randomize file and non-randomize file or non-randomize study.

At the hospital level, the (main) of patients in 2011 having had an NIHSS stroke scale administered, it was approximately 58 percent compared to 53 percent in 2010.

The (inaudible) was consistently more likely to be administered to the NIHSS although I don’t know why. Interesting to me because of my interest in transparency as these measures is not yet used in any public reporting – in a public reporting which is I don't understand

why. So my one comment is why isn't it useful for public usage. And I would hope that it would become so – but that's my summary.

Karen Johnson: OK. Thank you. As we look at the preliminary evaluations that you both did to this measure and I hope you can see that up on the screen. It looks like for important to measure and report, we didn't see a lot of problems with impact for performance gap that there is some question about the grading of the evidence looks like. So, maybe, to get us started with these things maybe one of the work group members would like to comment about maybe what they thought might be the problem with the evidence maybe.

David Knowlton: Well, so this is David Knowlton but somebody want to jump in. I want to create a dialogue that I think that the evidence issue maybe that it becomes – that haven't been a lot of extensive studies outside the original body of evidence on this particular measure comes from a single study (inaudible) measure and I don't know (inaudible) evidence.

Ramon Bautista: This is Ramon Bautista also adding to that statement. I mean, aside from the fact it's used mainly for tPA administration. There's not much evidence being useful for very many things at this point at least in a clinical realm. I mean, as we know, the NIH Stroke – the NIH Stroke Scale has its own limitation as well even as a stroke predictor, you know, it misses a lot of brain stem strokes for example.

Yes, but on the whole I think it's a good measure though. I think there are very many things that we don't know about stroke and I think having a standard way of trying to record stroke information as that is the first step in the right direction.

My main concern though is as we'll see later on is that by being – it's being administered by non-neurologist for example, ER doctors being trained to do NIH stroke scale not – they don't do it in a consistent way and trying to have this as a measure that is accepted nationwide for example because problems with reliability because not everybody probably know how to do that the same way.

Karen Johnson: OK. Anybody else has any other comments on anything under importance either impact or the performance gap for the evidence?

Risha Gidwani: This is Risha Gidwani. I am not a clinician so my concern from this was in terms of the link to the health outcome; it seemed from the way that this measure was specified that knowing an NIHSS for was going to improve patient outcome. So, I think really got a connection that is modified by in an intermediary step of having NIHSS for a change treatment practice that would then improve patient outcome.

So for the clinicians in the room, is that generally what you use as your decision point to treat with tPA or not?

Ramon Bautista: Yes, is there – and let me – among a whole barrage of different exclusion, inclusion criteria for but – yes, then the a stroke scale actually at this point in time helps determine the use of IV tPA into arterial tPA even the use of MRSA retrieval device. So there is – so NIH stroke is a tool being used for many different things nowadays.

Risha Gidwani: OK. And that would be considered a necessary piece of information before deciding to treat?

Ramon Bautista: That is correct. More like I said, one of the problem though is that in this day and age, it's actually used by merely by neurologist and if we don't have a neurology group where you practice I'm not sure if it can be as relied upon.

Risha Gidwani: So then I guess in terms of the outcome there's impact that it has in that outcome as Dr. Bautista is saying in terms of treating then there's a potential use what the developer's alluded to of the use of the stroke scale and risk adjustment which I think would be a fantastic use of this tool and they do allude to it and we can talk about this later when we get into usability.

I just wonder how that information is actually going to be captured not just in terms of the reliability that Dr. Bautista accurately points out but in terms of the collection and aggregation as reporting of the NIHSS score as well. We're going into an electronic medical record; does this mean that the data can actually be pulled out appropriately.

Female: And this is (Inaudible), can I say something about the importance of NIHSS?

David Knowlton: Sure.

Female: Yes. So, the AAN has written several public comments both CMS as well as the contractors looking at the next couple of measures. And we strongly believe that the strongest predictor of short-term outcomes among stroke patients is baseline stroke severity and that's what NIHSS collects. It has more predictive power than any other baseline variables so it is a very important variable to consider. And if you can't – if you don't collect it, then it isn't in these models to correct for stroke severity.

So, we would suggest highly that it is a very important piece of information to gather and to incorporate in any risk adjustment model, in particular the three that will be discussed later. And I actually like to be able to provide the letter that we gave to CMS on August 13th, 2010, that's in public record to this workgroup to understand our concerns surrounding predictive power of NIH Stroke Scale. Thank you.

Ramon Bautista: This is Ramon Bautista again. Just to comment on that statement. I fully agree with you in many ways to look at this as the – in the mental status exam, we now do for patients with dementia and it's come to a point where we're able to actually quantify the degree of dementia people have. And just hearing a number for example, you're getting – you can appreciate how bad they are and what kind of medications to prescribe for them.

And I agree that's having a way of standardizing stroke assessment and rather than just describing what you see in a stroke patient on your note is a first step in trying to get to that direction of being able to predict what kind of treatment to more accurately provide.

So I agree fully with that statement that there's actually – that the first step is actually being able to standardize the way we measure stroke severity and then go from there.

Female: And there are standard ways to train on this. We do also advocate an unbiased rather than independent. We think that the neurologist is more unbiased in the use of the NIHSS stroke

scale not that they have to do it but that they train and oversee because of the unbiased situation on neurology side versus (CAT) lab for example. So those two points would like to go on record. Thank you.

Karen Johnson: We need to kind of keep moving through on the discussion of this measure if there's nothing else from the steering committee on the importance criteria perhaps we can move on to your thoughts on reliability and validity testing of this measure. (Dave), did you want to start off and then the rest of the committee can jump in?

(David Knowlton): Let me go back to the information of reliability. The only comment that I saw on reliability that was not taken to account situations when the neurologist has the performance. It wasn't this (inaudible) reliability but I think what the study has indicated that clean nurses could perform reliably but I don't know if there's any other studies that looked to the reliability issues across raters.

The validity question I think is very silly within the acceptable range from what I've been able to see in the evidence from what was reported. It's the inter-rater reliability that remains the issue, I think.

Male: So they're – beyond this grade (inaudible) beyond the report, there have been extensive studies that have actually looked at this and a variety of operators so physician and nurses and study coordinators and even chart of destruction and shown good inter-observer reliability so we could provide additional support of data for you, if that's important.

Ramon Bautista: Will that include non-neurologist as well like ER doctors for example or, you know, family practice doctors doing the same testing – that would be shown for that group as well.

Male: A distinction for non-neurologist physician's predominantly emergency medicine physicians and nurses.

Ramon Bautista: But there's other inter – inter-rater reliability compared to neurologist.

Male: Yes, correct, compared to neurologist.

Ramon Bautista: Right.

Karen Johnson: So that would be great. Probably, Suzanne, would it be better to open their forums so they add it directly to the forum?

Suzanne Theberge: Yes.

Karen Johnson: OK. So, we will be doing that after the call so that you can add in that extra information about reliability. Thank you.

(David Knowlton): The results that this is – again, there was also the concern rate which most of you saw that the validity data that was presented to patients 55 years and older about this specification (inaudible) so …

Male: Right. I could address that so it is true for the 30-day mortality that that was the case but for in hospital mortality, we have paper that was cited that included adults having stroke from age 18 and above and once again shown to the greatest predictors as far as discrimination for in hospital mortality trumping all other clinical variables combined so that is one of the citations.

And importantly, beyond just mortality for functional status outcome of all types of Asians with acute ischemic stroke has been shown that their single most reliable predictor for functional outcome and then as was mentioned previously that their critically important role for determining the eligibility for, you know, the single, most effective strokes therapy of tPA.

Lee Schwamm: It's Lee Schwamm from AHA and Mass General. I just wanted to add that we didn't get what the guideline which is the resource as much of this data that we're presenting is contained. Many of those sites have non-neurologist performing and documenting the NIH Stroke Scale for us. And so, the model that was constructed within get with the guidelines incorporate all this different providers who are documenting NIH Stroke Scale for us. So at least an indirect evidence that even when documented by non-neurologist score that is recorded is a highly predictive of a short-term clinical outcome.

Karen Johnson: Great. Is anybody else have any other thoughts of comments about either reliability or validity?

Female: The statement from the AAN about neurologist performing is not necessarily that they have to perform that NIHSS. It's just highly recommended it because there are more unbiased that some of the training programs they advocated overseen by stroke neurologist. That's the point not that only stroke neurologist can do it. Yes, there's inter-rater reliability that shows anybody can do it. It's just the rule was around it are much more strict under the neurology unbiased assessment. Thank you.

Lee Schwamm: Is it possible – it's Dr. Schwamm again, I just wanted to make one comment about the previous section that I think was important which is aside from selection of therapies like for thrombolytics, the NIH Stroke Scale also predicts the use of ICU resources, identifies patients at higher risk of aspiration pneumonia and other complications including DVT/PE.

So, it is true that the link to outcome is through intermediate processes but having a marker of stroke severity I think greatly influences resource intensive care delivery and at times also provides an opportunity for discussions about ultimately end of life care. So, I think it exerts at dissecting a variety of different venues not just for thrombolytic therapy. Thanks.

(Karen): This is (Karen) from NQF. According to your (life) comment, I'm curious, I didn't see some of that in the evidence section. Did I just miss that in some of the citations that you mentioned, or are there the same listed in evidence section?

Lee Schwamm: I don't believe they are. I just was responding to one of the panel's comments about the link to outcomes.

(Karen): OK.

Ramon Bautista: This is Ramon Bautista. I agreed you. Unfortunately, you know, because we don't standardize the way we perform with it. We don't standardize the process of getting of these strokes geared consistently, we don't have a way of being able to use the data yet in a way of allocating resources. The only practical way of – at least from a clinical point of view is to predict tPA administration or not.

But I agree you that if we were to more standardly do this kind of NIH Stroke Scale on all stroke patients and there would come a point in time where we will be able to predict these resources more accurately.

Lee Schwamm: And I think that's exactly why having it available for public reporting in such an important step in the transparency and the ability to collect and record this data on all stroke patients.

Karen Johnson: OK. Let's go on to the next key sections, the usability and feasibility. So Dr. Knowlton, would you like to summarize this, too?

David Knowlton: I didn't hear what you said. I'm sorry.

Karen Johnson: Would you like to go ahead and summarize the evaluations on usability and feasibility for this measure?

David Knowlton: We didn't get especially in high school as a usability or feasibility but they're in moderate areas. The brief one that the (inaudible) is my concern is the inclusion of those reporting is possible but no explanation of why it's not considered to that. They certainly feel that (inaudible) and the comment is measured shouldn't be use for public usage and I question that.

Feasibility, if you have to be – as referred throughout this call, pretty easy to administer and pretty easy to record then the inter-rater reliability was addressed on the call. So, I think that any acceptable range for both (inaudible).

Ramon Bautista: This is Ramon Bautista. I'm just going to comment on that. I actually made that statement on a question in public usage. And I guess, I was thinking in terms of an individual family for an individual group of folks to find out that their loved one, for example, the stroke scale of 8 or 12 or 15 and I'm not exactly sure what to tell them what that exactly means because like I said that stroke scale is not a perfect science and there's (inaudible) kinds of strokes as well.

So, it's a bit maybe subject to misinterpretation sometimes by folks to might that understand why our love one, for example, did not tPA or did get tPA in that sense.

David Knowlton: Except that – this is David again, the issue of public reporting is mainly a reporting of the measure itself. So, the one percentage of patients 18 or over with an ischemic or unspecified stroke that received the NIHSS. The question and we know now is about 2/3, a little bit lower more recently in the 50 percent.

So, it's more in evaluation of the care that was given whether this particular measure was applied and why would the reluctance to publicly report whether this measure was complied with the (inaudible) complied with. That's what it means.

Ramon Bautista: Also because remember the NIH Stroke Scale just one of many criteria that we need to have before you actually give tPA. Example, somebody coming in at 3 ½ hours after stroke onset, you know, even if they may qualify from an NIH stroke score point of you may not necessarily qualify because there's some other exclusion criteria so take it as an individual measure by itself and not try to point out that you know what, you can qualify because you had diabetes or because your sugar was too high or something like that would be misleading to the patient.

So I think that by itself it may not be useful unless they can put context with other exclusion/inclusion criteria for tPA administration, for example.

Reva Winkler: This is Reva. Just to tell you that another one of the workgroups we'll be discussing very specific measures of tPA administration, so when the entire committee meets later this month, you will be able to see the full spectrum of the measures on the table and how the group of them will interact. Right now we're pretty much just looking each one sort of independently but we definitely have the opportunity to look at all the measures of the group.

David Knowlton: And piggybacking on Reva's comment, Dave Knowlton again, answers the doctor's point. As the score – there are certain patients who will be excluded from a (inaudible) because of the inappropriateness of them receiving tPA then they should be excluded and that would not impact on whether they shouldn't have had the state, the stroke scale administered to them. (Inaudible) to point now as fine as it is but it is – was a concern when I went to the scoring. Anyway …

Reva Winkler: So can you clarify again the key point I think you made which is this is not publicly reporting individual NIH Stroke Scale this is reporting the percentage of patients with acute ischemic stroke where stroke severity was documented.

David Knowlton: That's exactly – exactly …

Reva Winkler: (Inaudible) so that's critically important and we've discussed all of those rules why it's not use for public reporting yet as we don't yet have NQF endorsement but it is and get with the guideline stroke reported back to hospitals and used for quality feedback and we believe it's been useful in that regard.

David Knowlton: From where I sit where we get public reporting, we had data compliance and its compliance is what we're looking for. But we will take into a context of everything else later on. I'm concern about our time I think of probably key areas that you want to address.

Karen Johnson: Right. So anything from anybody else on that measure before we move on to the next one?

David Knowlton: The conclusion of this measure was to adopt the preliminary consideration and I certainly agree with that.

Reva Winkler: Sure.

Karen Johnson: So let's go ahead to the next measure then on our left, Measure 2017 Stroke and Stroke Rehabilitation CT or Magnetic Resonance Imaging Report. And Dr. Bautista, this is yours.

Ramon Bautista: OK. Thank you. So, for this particular measure, the numerator was the final report of initial CT Scan or MRI included documentation of the presence or absence of each of the following – hemorrhage, mass lesion infarction and the denominator was all final reports for CT scan and MRI perform within 24 hours of admission to the hospital or in an out-patient imaging standard to confirm that the diagnosis, stroke, TIA or intracranial hemorrhage for patients more than 18 years and older. So that's the measure here.

So, just to give my thoughts on this, when you go to the (inaudible) in the notes here on the impact in opportunity for improvement, I mean, it's probably stating the obvious that we know that a substantial percentage of patients with stroke will come in with stroke like symptoms in fact have (inaudible) something we all know.

And the next day they do – they do rather some performance gap in 1b.2 and I guess at least for the sponsors if they just try to help me understand what they actually mean by performance, what is the actual performance gap here? Is it the fact that (A), not in a CT Scan or MRI were done or (B), that the CT Scan or MRIs do not document the preference or absence of what they want to measure here. I can get some clarification on that at this point.

(David Seymour): This is (David Seymour) from ACR and PCPI. The gap that we observe is in the reporting of the nines exclusion criteria for acute stroke patients which hinders the prompt decision-making and correct decision-making with respect to tPA administration in this population and, of course, any subsequent management decision.

Oftentimes, just to give an example or recently published paper by (Prevedello). I don't know if that's been circulated among the workgroup reported that only in 58.1 percent of cases where the three nine elements reported. So and this was just published in the last month so this is quite current data.

Ramon Bautista: So, let me just clarify it again. So, other than that, what elements where these again that were not – were reported or not reported – I'm sorry, I'm just missing something here.

(David Seymour): Oh, sure. The stroke hemorrhage mass effect, acute infarct hemorrhage mass effect.

Ramon Bautista: So in other words (inaudible) whether or not they report actually contain the statement that said there is or there's no stroke or bleed or mass or tumor. Is that pretty much what a …

(David Seymour): Pretty much, yes.

Ramon Bautista: OK. OK. I understand that. Thanks for clarifying that. Let me just give a few – just some thoughts here though as far as the impact is – because I was standing myself and again, I'm not a radiologist but I think we have somebody from the (inaudible) radiology here.

I would imagine that when a radiologist would took an MRI or a CT Scan, they would – they would make a mention whether or not they saw it would be obvious to them that there was a stroke or a bleed or something structural going on. They would make a mention of that. Is that not part of the standard reporting done by radiologist?

I mean, I'm read EEGs, and I pretty much would say that. There's a seizure activity of – if a see seizure activity so I'm just trying to understand if there are a deficiency in the way radiologist dictate their report that don't even mention bleeds or strokes or masses or tumors?

(David Seymour): Well, if I'm a radiologist then perhaps it's a bit embarrassing to say this but sometimes these exclusion criteria aren't mentioned explicitly.

Ramon Bautista: OK.

(David Seymour): Even though when it's actually quite interesting what happens when these findings are present or absent. When an infarct is present then adherence to the other pertinent negatives is quite – has a higher odds ratio I can't recall off the top of my head what it was but it was statistically significant.

On the other hand, in the presence of hemorrhage, the other nines exclusions aren't mentioned at all with again a statistically significant decrease in odds ratio. So, sometimes a radiologist simply says normal scan or no acute abnormality seen or something like that leaves them the element of doubt in the …

Ramon Bautista: So it's basically, I mean, that – I'm sorry, I want to be blunt here, it's basically a misread. I mean to see that there's no abnormality found and there's a stroke or (inaudible).

(David Seymour): Oh no, this measure does not address the – regrettably, we're not to the point yet of addressing the correctness of the interpretation. We're simply addressing the inclusion in the report of the pertinent information that's required to make an expeditious decision.

Ramon Bautista: OK. I understand. Thanks for clarifying that. I was ensuring if that would – so having said that and then agreeing with you that because of the way things or probably need to specifically mention those statements over there. The second question is I mean how good really is an MRI or a CT Scan in speaking of these abnormalities we want to pick up.

So for example, I mean, we know, for example, that many lacunar infarcts may not be picked up on a CT Scan. A TIA may not be seen on a CT Scan or an MRI for that matter. So, in other words, the negative predictive value, for example, of a CT Scan in speaking of a stroke or an acute infarct much of a lacunar infarct is not all that great anyway.

So, for us to require, for example, that you document this thing and make these statements aren't this sort of like putting yourself in a bad shape where they're going to say no acute infarction was seen and you want that to be a measure of sums of quality whereas in fact the test is not even good enough to really pick up.

(David Seymour): OK. That's an interesting question. I think that whenever one interprets a diagnostic test as a radiologist for example or a neurologist interpreting an EEG. And then when you're the

receiver of that information as let's say the neurologist or ER doctor in the case of an imaging report. You have to take into account the limits of the method. And I believe that when a – I believe that when the decision to administer tPA is made, the criteria respect the limit of the method yet still show a days statistically significant improvement in patient outcome. So, yes, I mean, we know that we're not pathologist, for example.

Lee Schwamm: This is Dr. Schwamm. I'm a stroke neurologist. I just wanted to comment from the sideline here. I think the issue might be better reframed in that. The clinicians who are not radiologist and oftentimes are not neurologist are looking for specific language to help them understand whether or not the scan can be consistent with the treatment of thrombolytic therapy.

And if they don't have a very highly technical knowledge or radiology terminology, they might not recognized synonyms or other reporting that would infer the scan was safe for treatment but in fact not have specific exclusions listed in the report. Is that the issue?

(David Seymour): Very well put, Dr. Schwamm. Thank you very much.

Lee Schwamm: So it's really about creating a standard format for reporting so that non-radiologist can know with certainty that a scan is safe to treat with a thrombolytic.

Ramon Bautista: And I'm looking at this few, maybe a couple of steps out there trying to explain to a family, for example, in a litigation case for example. You know, if this become standard, trying to explain to them that the report did not say there was a stroke yet you gave tPA. What's going on here, you know that sort of thinking.

Just to my mind these makes the measures more difficult to use out there for public consumption. There are there are least limitations out there.

(David Seymour): Let me try to address that one. If this were used for public reporting that would be in a manner I believe similar to the NIH scale and it would be a performance measure for radiologist based upon their provision – necessary information for stroke treatment. With respect to a litigation situation regarding the decision to administer tPA or not, I believe that the guidelines for tPA administration prefer a negative CT Scan.

So, I believe that the – and Dr. Schwamm, I believe would be more expert at that type of thing than I would but the finding of the CT Scan with, you know, an infarct below a certain size as the inclusion criteria.

Karen Johnson: OK. In the interest of time, I notice that to (inaudible) this preliminary evaluation had (no) for evidence maybe we should talk a little bit about evidence.

Ramon Bautista: Well, I went over the – I look at the date here, says quantity, quality and consistent of the results across cross study. And what I'm reading here – I'll read the last line that the ACR guideline does not address the overall quantity of studies and the body of evidence nor the overall quality of study, nor the consistency of results across study.

So, I guess, I don't have much information here from the data here that this measure have good evidence based from the measure I'm reading over here.

Reva Winkler: So sounds like your concern is based more on the lack of information or insufficient information which again does not allow the measure to pass that sub-criteria.

Ramon Bautista: That's right. It's not that it's a bad measure but I don't see much in terms of evidence for the measure.

Reva Winkler: Right.

Male: There's some more concern (inaudible).

Reva Winkler: OK. Go ahead.

Deidra Joseph: Hi. This is Deidra from the AMA-PCPI. I just want to address that really quickly. So, we have several guidelines to support the measure. However some guidelines or guideline specialties decided to create the guidelines not include quantity, quality and consistency data within the guideline. So, there was no way that we can include that. Additionally, I did add just very recently one systematic review regarding the quality of written radiology report as well as an additional study. I just added that probably maybe half-hour before the call.

So, there's not, I mean, unless we could obtain a technical report from the specialties societies that publish the guidelines. We don't have access to the literature that they reviewed in order to develop their recommendations although the grade that they use to grade the recommendations do speak to those particular things which is also included in the evidence section.

Reva Winkler: This is Reva. Just in general, the criteria do speak directly to the studies and the evidence and guidelines should be based in evidence but we aren't – we want – are really asking you to go beyond the guidelines and get to the underlying studies and evidence. And whatever literature of use and necessary to support that but that really is the information that is necessary for the committee to apply the criteria.

So, it sounds like there are some additional information that's been added that the committee will have in for the meeting but I think this is an important point that we really are looking in an evaluation of the study as described in the criteria.

Male: Please go ahead.

Female: So, part of this is a structural process issues so the documenting and specific findings of CT or MRI reports are really important with symptoms or diagnosis. And so, part of the problem is the real specific problem is documenting absence of presence of any of those three. And so, it's a process structural type outcome and it does improve understanding of patient or improve outcomes if it is documented in these reports and specifically its presence or absence of each of the following – hemorrhage, mass lesion and acute infarction. Those are all – the upper hand and that they typically don't document all of the – or presence or absence of each one of those three.

And so, it's a process outcome structural issue which I spoke into I think in 1C1 and the importance of how it – the symptoms or the diagnoses impact improved outcomes committee.

Risha Gidwani: This is Risha Gidwani. Reva, thank you for your comment. I feel in a similar manner. In terms of 1C1, it says that the process of documenting either the signing is link to improve health outcome but it seems to me that again, there's a missing link here and said the process of documenting and declining will then lead to improve healthcare processes and those processes themselves maybe link to improve health outcome.

So, seeing actually the literature supporting that link from the documentation to the improve processes, to the improved outcome, I think would be very beneficial for us as a committee.

In terms of the guidelines and essence that was given, it would also be very helpful as prior to the next meeting. We were actually able to see what specifically each guideline said. I see here that we have numbers of guidelines but we don't have what those guidelines actually referring to.

So instead of a guideline number one, it would be great to have guideline number one and then what verbatim that guideline states and then what the quality the evidence is associated with that guideline. All right. It's not …

Male: I agree.

Risha Gidwani: Also, in terms of 1C.12, the grading scale used for the body of evidence, there are actually two definitions provided at each four level A, B, and C. So, it's unclear whether level A indicates data derived from multiple, randomize clinical trials or data derived from multiple prospective cohort study, the use the reference standard applied by mass evaluator. Those are actually pretty different in terms of the quality of evidence. So, some more specifications would be (inaudible).

(David Seymour): And the question here is not on – if I may clarify just so we can provide you with the information that you require prior to the next meeting. The question is whether the reporting of the specific data element is tied to outcome or whether the performance of imaging itself is tied to outcome. Which of those two questions, or is it both?

Reva Winkler: It's really the focus of the measure. So if you are measuring the reporting of specific elements in your report then that's the body of evidence that should be referenced.

(David Seymour): OK. Thank you.

Ramon Bautista: This is Ramon Bautista again. So just to go back to your performance gaps, (inaudible), are you saying that currently radiologist do not consistently report the presence or absence of stroke, tumors, or bleeds in patients who present with stroke-like symptoms and such absence of reporting actually has harmed patients. Is that something that is known from a radiology point of view?

(David Seymour): The first part, yes. The second part, I don't believe has been documented but I believe it considerably a valid inference.

Ramon Bautista: I mean, because if that's the case and that's pretty obvious, we need to have something of that measure, I meet this from a (making) point of view. If that's what's been happening, then obviously we need to do something about that.

Female: Yes. It is critical …

(David Seymour): Well, that's the hypothesis. If people aren't getting the clinical information that's required for clinical decision-making then perhaps clinical decision-making is hampered and therefore patient outcome is hampered.

Female: Yes. And the other critical word is final reports because there could be multiple CT MRIs and yet and some of those document one absence or presence of one of those three and yet the final report does not repeat those findings even though maybe they looked back at some of that.

So anyway, it's the terms in the measure statements that are critical to understand final reporting of these – yes, thank you.

Karen Johnson: Thank you. We are fast running out of time for this measure. So, we would like to keep our comments if we could just to the steering committee members and the developers but as we go on, scientific acceptability of this measure most like there was some disagreement on reliability and validity. Most of the members have any comments that they want to bring out on either of those issues.

Ramon Bautista: Well, reliability testing was performed in just three sites (supporting) 95 patients. So, that's a small number compared to I mean the other measures we're measuring out over here. And that's just one thing – one thing to think about.

(David Seymour): The recently published article I just reference covered 4,000 and did show sufficient reliability that automated data extraction was possible using a natural language algorithm. So, I think that's time safety evidence of reliability and in terms of consistency.

Deidra Joseph: (Dr. Seymour) and this is Deidra. That data have also have been added to those form.

(David Seymour): Oh, good. OK. Thank you.

Karen Johnson: OK. Any other comments about reliability or validity as a measure?

Ramon Bautista: (Inaudible) to my mind again my first statement about the limitation so far of an MRI and a CT Scan in determining strokes or, you know, or acute infarcts or TIA for that matter is an issue. So even though – so the statement that there is no evidence of a stroke here in the patient with a TIA and, you know, this would make a whole lot of sense to me or so – I think all the intent of the measure is good but please measure is noted that there is a stroke or a bleed or a tumor or not. How bad is the (inaudible) how valid that is and how much that correlates with the clinical scenario is not 100 percent – is not straight.

Deidra Joseph: So, just for clarification. This is Deidra at AMA. The evidence that you're asking for, is it to support (inaudible).

Ramon Bautista: Go ahead. Sorry.

Deidra Joseph: I'm sorry, is it to support CT and MRI for diagnoses of stroke? Is that the evidence that you're asking for?

Ramon Bautista: Yes, if you want to – if you're doing this measure in order to state that there is a stroke or an absence of a stroke or a bleed or a brain tumor for that matter, you know, the assumption is that this would correlate clinically and help the clinician make the decision whether or not to treat with tPA or some other procedure on the patient.

But if you know that the measure by itself – a CT Scan by itself is not perfect when it comes to picking up TIAs or even acute stroke then it's not a valid test. It's not a valid measure basically unless you put down there and normal CT Scan is not relative to stroke and then it was like a double negative there, right?

Karen Johnson: OK. Does anybody have any comments and again, we are kind of tight on time for this measure but any comments about either usability or feasibility.

No?

Female: I just want to say that CT and MRI are more common than any other and some facilities don't have a multi model approach using such us the imaging things that are available. So CT and MRI was the most common and that's they're called out.

Risha Gidwani: Can you give us background about how this data would actually be collected?

David Knowlton: I saw the note, the unintended consequences – do we know what that implies.

Ramon Bautista: So, I mentioned that myself. In other words, my concern is – and as I mentioned earlier, is we have a situation where somebody comes in and there's a perfection by, for example, family that the MRI or the CT Scan, you know, did not reflect the level of care and there's going to be a lot discussion of – you know, because of the test by themselves are limited basically.

So the question is – so the question is either you have a measure here that actually mandates reporting those things but they could be misinterpreted. I think they could misinterpreted by patients, by families, by lawyers.

David Knowlton: OK. So it's more liabilities …

Ramon Bautista: And a usability issue as well.

(David Knowlton): OK.

Reva Winkler: Also, can you developers address the collection of data elements and how that is expected to (inaudible)?

(Carrie Christiansen): Hi. This is (Carrie Christiansen) from the AMA-PCPI. And we presented information from our two different projects in our (inaudible) materials that we provided you. One of them as Deidra mentioned was just added today when the forum has reopened for us. But the first was a measure testing project using a very standard measure as being protocol and you can see in there that the Kaplan scores came out extremely high indicating, very high reliability.

And then the second measure testing project was the one on the – was mentioned before the (Prevedello) study that was just published recently in the Journal of American College of Radiology and that also showed extremely high agreement rate for the measure, data element that was at the data element level. And those are showing agreement rate all well above 90 percent.

Ramon Bautista: And let me just ask you a question. What are you actually measuring there? Reliability of what?

Karen Johnson: (Inaudible).

(Carrie Christiansen): So the first – the first testing project actually was looking at three different site and doing inter-rater reliability so having one human abstractor and another human abstractor, look at the patient and determine if they met the measure or did not meet the measure and the agreement percentage is there as well as the Kaplan scores were almost perfect for the numerator, denominator and the overall measure.

And then the second project was actually a comparison between a human manual abstractor doing visual inspection of its (inaudible) versus an electronic health record automatically reporting the measure. And again, the agreement percentages were very, very high at the data on its level.

Female: (Inaudible).

Ramon Bautista: So (inaudible). So if somebody actually had a stroke, came in with a stroke that it would appear to be an ischemic stroke at least in clinical presentation and the CT Scan showed a stroke. Is that where you're seeing it's a reliable measure? Is it among different people reading the CT Scan?

(Carrie Christiansen): So, reliability of a measure. Reliability and validity of a measure, the way that NQF criteria explain it and sees interest, feel free to jump in and explain it if I don't use quite the right words. It's checking to make sure that the measures able to produce consistent result.

Ramon Bautista: Right. And I understand that part but are you seeing that, for example, if a CT Scan showed a left MCA stroke, it would be a reliable measure of three different radiologists independently looked at the CT Scan and said, "Yes, there is a left MCA stroke." Is that what you're, I think, reliability?

(Carrie Christiansen): No, it doesn't have anything to do with the result, the actual clinical results of the patient. What has to do with is one person looking at the patient's chart and determining if they met this measure or not and another person looking at the patient's chart and determining if they met this measure or not and comparing those two results to see if those two match. That's the reliability of the measure itself.

Reva Winkler: It sounds like you're looking at the reliability of whether people can determine whether the documentation exist regardless of what that documentation (inaudible).

Ramon Bautista: Is that what it is? This document is …

Female: Correct.

Reva Winkler: And that speaks to reliability and that's fine. But I'm asking about the feasibility which is I'm wondering how exactly are these data element suppose to be captured. Are you suggesting that once NQF (inaudible) this measure that all providers should be engaging in chart with you to assess whether this reporting has occurred?

Judy Burleson: This is Judy Burleson at the American College of Radiology. Just as an example, the measure has been included in the PQRS program since 2007 so it was reported through use of CPT two codes on claims and sometimes in registries. And so there is a code that indicates that the data elements were included in the report as asked by the measure.

So you include a CPT II code on your claim that says, yes, I address presence or absences, image, mass lesion or acute infarct and statistics from CMS for the 2010 reporting year – a little over a half of physicians that reported that measure reported it correctly. So it has been in place and used for at least four years.

Reva Winkler: And so this is – I noticed that the measure is on both the in-patient and the out-patient side so the CPT data are also able to be collected on the in-patient side?

Judy Burleson: Yes. It's a physician reporting the claim for professional services. So it's a professional claim.

Reva Winkler: OK. All right. Thank you.

Ramon Bautista: This is Ramon Bautista again. You know, at the end of the day, I think what you really want to find about whether or not by having a radiologist explicitly say, yes, there is a stroke of brain tumor, a bleed or none that somehow that correlates with better patient care at the end of the day. I think that's really what this measure should try to focus on.

Karen Johnson: OK. That's really great discussion. And I hate to cut it off but I really need to because we still have our outcome measures to talk about. And they are even a little bit more complex. So, let's go on and go to the Acute Stroke Mortality Rate 0467 to AHRQ measure. And, (Carrie), I think this is your measure, right?

(Carrie Christiansen): Yes, it is. So this is the in-hospital acute stroke mortality measure is the only endorsed measure we looked at. It was endorsed in 2008. So, it really, is a maintenance review. It

looks at the (size) of discharges for in-hospital death on one patient for the sample size and it's also called the stroke. So it's the number of the deaths overall discharges for 18 year and older, 18 years of age and older for diagnosis of stroke.

There are some exclusions. The one that is transferred to another short-term or acute care hospital, one is MDC, Medical Diagnosis Group, (Inaudible), child birth, and care (inaudible). And the other is that their team has been discharged disposition data, gender, age, some of the key data.

So it's an outcome measure that you just administrated, claims data to – there's the story for this data. The source – this wasn't it. These were interesting in this group and that so one area we all agreed on what's on impact. But that's the only area we all agreed on and that's preliminary assessment.

So clearly, some from an impact perspective, there's a lot of strokes over here with the mortality rate of about 17 percent, greatest mortality rate occurs on 30 days, approximately some days in 2008, approximately 46 percent of those deaths occur in hospital and this is an in-hospital measure.

And the group all agreed in terms of the threshold be met on this. The outcome measure, there was a linkage to make between structure, for example, strokes on a process, for example, if there's any complication, an outcome of in-hospital mortality. And because it was a maintenance measure, there was an expectation of the position of data using this measure demonstrating – oh, gosh – performance staff or disparity and those data are presented as an (HCAP) data but 2009 national inpatient sample.

And then there are variation shown in risk adjusted mortality for hospital type, hospital sizes, region or country where the hospital is located, size of the – I'm going to say, study, for lack of a better words, of where the hospital is located, the bed size of the hospital, the income level. So that's the first criteria.

Karen Johnson: OK. As I look at this preliminary evaluation results here, it looks like most people were pretty happy with impact and performance gap.

(Carrie Christiansen): Right.

Karen Johnson: Does any – are there any comments that we need to make on that or should we just go straight into the …

(Carrie Christiansen): Unless somebody else wants something here.

Karen Johnson: OK. Let's go on to the scientific acceptability. I think that's going to be the meat of this particular measure.

(Carrie Christiansen): I think so, too. So scientific acceptability and some of the reliability and the validity, there was typically (named) on both of those criteria. Let's see.

So, in-hospital mortality, reliability data are presumptuous, which looks as if to me as solid. I think for my question on that was, if you are an assistant where you could easily look patient throughout of the acute care setting, to die somewhere else is that they would not be captured and (inaudible) identify that as a possibility.

Let's see. Should we just talk about reliability because somebody, it might have been (you), but I'm not sure, have a lot of comments on reliability for this. Do you want to jump in and talk about reliability here?

Risha Gidwani: Let's see. Actually, mine was not the reliability, because mine was the …

(Carrie Christiansen): Oh, OK.

Risha Gidwani: Yes, actually …

(Carrie Christiansen): Somebody had a lot of comments of reliability, so if anybody wants to – I first write down reliability. I didn't have any major questions on this other than the movement to a non-acute care setting. Are there other issues on reliability that people would like to bring up?

I see a hand out there. I may – has it changed? Let me see. Are you looking at an old one because I'm looking at the newer one, the newer version. And I think, Suzanne, you just sent it out?

Risha Gidwani: Yes. You know what, actually, but yes, in that one. I was just looking at the computer screen here in the webinar and I think that is the old one. So, yes, on the newer graph that was just circulated, I did have a number of questions about the scientific rationale, but I read the methods document, the Quality Indicator Empirical Methods.

And it actually – this is the document that is applicable to all of the HRQ measures and so I was hoping that we can get some specific information actually for stroke which would allow us to assess the specification to that model that, for example, we don't even understand.

We have no documentation of what variable which are included in the model with their beta coefficients are. And that would be important for us to assess especially in light to the fact that we can have another measure about mortality and we want to make sure that we are able to harmonize or assess them both properly.

Karen Johnson: Risha, this is Karen. I don't know if you were able to see a document, the name of the – called Risk Adjustment Coefficients for the IQI version 4.4. That one just have the parameters with some data estimates and key values.

Risha Gidwani: OK. I will have to take a look at that.

(Carrie Christiansen): On page 11.

Risha Gidwani: Page 11.

Female: Yes. There's a – there's a table for each of the different IQI. You're right. It's a document that contains information about many measure, but within it is specific information for this particular measure. Table Nine.

Risha Gidwani: OK. I want to take a look at that. So there were also a number of specific questions that I had in there and that – it actually probably is not the best used to the entire committee's time to go through this one by one but if there's some way that we could get some feedback on that …

Female: Sure.

Risha Gidwani: … certainly in (business) format that would be really useful.

Female: OK. What we can do is pull this together and then forward it to the measure developer. (John), are you still with us on the call?

(John): Yes.

Female: So, well, you can expect to receive something from us as we'll forward Risha's comment.

(John): Yes.

Female: Thanks. Anything else?

(Carrie Christiansen): And this is – I'm sorry, this is (Carrie) again. I have a question on the validity as well. As I said that – as (inaudible) insufficient. So comparing chart data with the administrative claims record and the accessibility and specificity which looks good with relatively tight confidence and (inaudible).

But it looks like that (somebody) want to focus on the denominator of the, you know, like, can I catch the stroke and am I doing a good job at that as opposed to, am I (clustering) death which is the numerator. Have I just missing something or take – in respond to like – just speak to that.

Male: Well, I think it was most – it was most likely Patrick Romano who wrote up tight the section of the form and he is unfortunately not on the call today. So I'll relay that question and we can have that answered in the full steering committee meeting.

(Carrie Christiansen): OK. That would be great because – that would be great.

Karen Johnson: (Carrie), this is Karen and you actually put your finger on one of the – one of the things that we had noticed here. We do say that if they do critical, if they do data element or whether they need to talk about the critical data element and as you say the mortality itself is a critical data element. So good call on that.

Female: Any other comments from the committee or questions that perhaps we can forward to the developers for them to be able to thoughtfully respond to and we can have that to you before the next – before the meeting?

Ramon Bautista: This is Ramon Bautista again. I actually – it's a good measure, ma'am. My only concern is that it's not risk adjusted so I imagine that the typical stroke patient across different medical centers may vary.

You know, you might be in the inner city area or maybe in urban setting and that they might have a different kind of stroke patient you see. So I believe signs of risk standardized or risk adjusted measure to make it more even playing field would be – would be a good step, otherwise, you'd be comparing apples and oranges across different centers, you know, when it might come to stroke here.

Male: Yes. It is duly noted, beginning on page 11 that lays out the risk adjustment for the measure. It is a risk adjusted measure.

Ramon Bautista: Right.

Female: Is it risk adjusted for NIHSS because that's one of the most important predictors for mortality? And in addition, it's a disservice to the public in a way to rank hospitals based on outcomes that aren't risk adjusted for stroke severity and you'd – is there are a lot of primary stroke and competent stroke centers that get a lot of the sicker patients.

And if they're ranked based on that without the adjustment, then this – the public will look at it and say, "Oh, my gosh, that's a terrible hospital and yet they get the more severe patients."

Male: That's not one of the covariates at this time. We used elements taken from the claim and that point to table nine, in that file of the covariates and coefficients. It does note to see statistic as 8.894 so it's there rather well.

Female: Yes.

Risha Gidwani: This is Risha Gidwani, and I could sort of those points here that, you know, the stroke severity at baseline is really important but just from a logistical point of view, that's not captured in the billing data. It's going to be difficult to actually include that because there's not going to be the data element to work with. It was the information rather than data that could be used in the model.

Female: I have the …

Risha Gidwani: I just have to look at – I did – I'm sorry, I did look at page 11, I found that section. I hadn't seen that before. And I do see that risk adjustment coefficient, I see that they – we have labels for the APR-DRGs. In the next round, would it be possible for you to actually note what – instead of number, the number of the APR-DRG, the actual name so that we can get a sense of whether some other measures of stroke severity or co-morbidities are included?

Male: I can ask (Rianne) if they will allow us to relay that, that's a good possibility. So that's a good point. We'll take that to her.

Risha Gidwani: Thank you.

Ramon Bautista: Hello. This is Ramon Bautista again. This measures one of those that can really make or break a stroke center and I'm going to say that in a less than correctly, you do have an awfully lot of people being unfairly discriminated upon for doing a poor job whereas in fact, what they have is a different profile of stroke patients that they see.

Female: OK.

Female: Yes. And this statistic argument is that NHS is alone and just alone just produces these statistics.

Ramon Bautista: Yes, I agree.

Female: (Inaudible).

Ramon Bautista: You better careful about what you add as a risk adjustment over here. You need to make sure it's comprehensive.

Reva Winkler: Right. So Dr. Bautista, just in terms of your comment about the measure being risk adjusted, are you clear now and can find the information about the risk adjustment methodology?

Ramon Bautista: Yes. I just want to know exactly what these different DRG measures are.

Reva Winkler: OK. So that's a fair question and we'll hopefully try and get that information for you.

In terms of moving things along, it could be that a couple more measures. Any other questions of the reliability validity information which we've got question that will go to the developer? How about usability or feasibility for this outcome measure?

(Carrie Christiansen): Usability, this is (Carrie), and feasibility, we were on three different cases. So usability, I actually thought it was good. It has been used – and is publicly available and (inaudible) our system although when I – when I went on and actually look at it, some of the states they didn't have data available for this measure and choose for the (inaudible) and to mark, here's the (IQI) for university health to enforce them and then another crew.

So from usability, to start, it will actually rank as high but somebody else acted as low and somebody looked at (inspection), I don't know, so we were not on the same page on this.

Ramon Bautista: This is Ramon Bautista. I'm going to probably withhold my judgment on that based on what the DRG – what the DRG measures although we're going to use to risk adjust.

Reva Winkler: OK. Other comments from committee members? I see (inaudible).

Risha Gidwani: Well, this is Risha Gidwani. I think it's really hard to assess the usability without sort of having a better understanding of the model itself and so …

Reva Winkler: OK.

Risha Gidwani: … from my perspective, I'm tabling – I was the one that said insufficient and it's not because it – just because I need more information before I comment.

Reva Winkler: OK. All right. So we'll try and get that for you to consider prior to the next meeting. Any issues around feasibility you want to bring up? This is a measure based on administrative data.

Male: (Inaudible).

Reva Winkler: OK. So we need to move on to the next measure which is again another stroke mortality measure. This time, it's the 30-day, all-cause, risk-standardized mortality rate. It is a new measure to NQF and this is brought to us from CMS and their friends at Yale Center for Outcomes Research and Evaluation. Who's measures is this?

Female: This is Risha's.

Reva Winkler: Risha, this is yours? If you want to…

Risha Gidwani: Yes. I just went on as well, yes. So one moment, here? OK. So this is a 30-day, all-cause mortality rate based on a hospital level. It's for all patients to have acute ischemic stroke for 65 years or older.

Patients who are excluded from this measure are those who don't have a 12-month claim history, those who were transferred from another acute care hospital, folks that had an inconsistent or unknown mortality status, who were discharged against medical advice, or those who were enrolled in a Medicare hospice program in the 12 months prior to initial hospitalization.

It's been outcomes measure. It is based off of administrative claims data. It is risk adjusted using a hierarchical logistical regression model that takes an account factors as well as hospital factors so that there maybe correlations in patient outcomes within hospital due to hospital factors.

The covariates that are used for risk adjustment are patient level factors including transfer from another emergency department, age, gender, and a variety of clinical characteristics. The covariates that were used for risk adjustment were selected based off of this statistical significant, the literature, and/or clinical judgment.

In terms of the importance to measure and report of the three folks who responded, all three felt that it was important to measure and report. In terms of the evidence as an outcome measure, everybody felt as though there was good evidence in terms of use of this measure. And then there was some discrepancies when we got into the specifics of the reliability and usability and feasibility.

Female: OK. So you want to go ahead and open it for any questions about reliability?

Risha Gidwani: So I had a – you can see that first paragraph up there and those were a number of questions that I had. I did receive this morning a document responding to a lot of those questions. I

would actually like to take some opportunities to engage with the developers and ask them follow up questions.

But before I do that, let me ask my colleagues on the steering committee if they'd like to add anything.

Female: No.

Male: No.

Risha Gidwani: In terms of sort of the – hey, take a big picture step back. This measure is looking at a predicted mortality versus an observed mortality. And I'm still struggling with the use of predicted mortality as a numerator rather than observed data. I did receive some feedback on this, but if the developers could take a moment to explain the rationale, that would be very useful.

Susannah Bernheim: Hi. Sure. This is Susannah Bernheim from the Yale team. I have a more challenging aspects of our measure that is similar to the other outcomes measures that we have brought to NQF and some of which are now currently close recorded.

Conceptually, this measure is using hierarchical modeling. The goal of the hierarchical modeling is to account to the fact the patients are (a cost) within hospital though it's already expected. It's very much analogous to an observed over expected. So you can think of it as being similar in one of the things that we pointed out in response to some questions we got is that the hospital's actual readmission data is use in coming up with the predicted.

What the predicted does is it allows us to compare how a hospital is performing with its given case mix as the predicted to how an average hospital would perform with that same case mix.

The observed over expected, one of the pitfalls of that is that if you have a hospital with a very low sample size, you have a much harder time trusting that observed right, so if you only have 10 stroke patients and six got readmitted one year, the next year, it could be three. You already know if they have a 60 percent or 3 percent – 30 percent mortality. So that's one of the problems that falls by the hierarchical model.

The other is this accounting for clustering and sort of appropriately taking account the structure of the data. So I mean, I can – we spend a fair amount of time discussing this and I can spend more time on it, but I know you guys have limited time.

Conceptually, it is very similar but it allows us primarily to account for case mix, sample size, and the structure of the data.

Risha Gidwani: Is there a thought – I mean, I certainly agree with the issue of the clustering of the patients within hospitals. Is there a thought though to using the hierarchical generalized linear model to create the denominator and then using the observed data to create the numerator in that way you're still allowing for that lack of independence of patients with in-hospital?

Susannah Bernheim: Well, you still have to have those sample size issue there, right? So your observed is really unreliable if you don't have a reasonable sample size and what this allows us to do is sort of take that into account as well.

Risha Gidwani: So I see that that's an issue. I just wonder – so I guess, we've got a trade off to make here and that's whether, you know, we sort of have the issue of some hospitals having very small sample sizes or the issue of all hospitals using predicted data.

And, you know, from my perspective, I feel like it's always, you know, a better call to use the actual data than the predicted data especially given the – I think, is the fifth – yes, the area under the ROC curve for the models .602.

So that indicates the model is not able to discriminate as well as one would hope between the patients that actually – or I'm sorry, we're at – we're at the mortality, not the readmissions, excuse me – that readmissions model that do – that do pretty well.

But it's still not a perfect discriminant ability between the patients that had readmission and those who don't, so we're bringing that lack of discriminant ability into the numerator when we use the predicted readmissions and we're bringing that in for all hospitals.

Susannah Bernheim: Right. So one thing – I just want to go back to an earlier statement I make because I do think this is confusing. The observed readmission rate is a strong contributor to the predicted, right? So it's not a – it's not distinct from the observed rate. I mean, the observed rate drive the predicted in a large part, but it then is, if you will, adjusted form, this is a little bit – I'm being a little bit measuring, but it is essentially then adjusted for case mix and sample size. I mean, that's essentially what's happening in the numerator.

So we're not divorcing ourselves at all from the hospital's observed rate. You know, because these measures are meant for a public reporting, our approach has always been to be very careful not to unduly identify out wires and the hierarchical modeling. I've done both the numerator and denominator, it, you know, allows us to be cautious in that way. So they were accounting for all of the factors that might be contributing to that observed.

But, again, it is – it is an – it is a version of the observed. It's not, you know, it's not sort of a simulation model done in some other setting. I mean, it is based on the actual readmission rate by that hospital or mortalities. Sorry, right now, I'm confusing which measure we're on.

Female: OK. But I think – I think we can actually – I mean, the approach is the same across each of these models. It's the data that we're using, the outcome experiment so if you please (inaudible) have the conversation closely.

Risha Gidwani: I just, you know, I see what you're saying and I understand that the – from reading about this report that I was sent by the CMS, the president of the statistical society, I see that. But we're still building in some level of uncertainty because we're predicting rather than using actual.

And so I wonder if potentially some of these issues with the small sample size that they might be handled better by some sort of smoothing estimator, for example. In terms of the case mix adjustment, if we use observed data, then we're already predicating it upon the case

mix adjustment of that particular hospital and then that denominator of the expected could actually just be for hospitals that have the same case mix as the numerator.

So, you know, this might be a larger statistical conversation and I'm happy to have it. You know, I just want to have a better understanding and maybe walk through potentially the decision point that you guys went through to reach the conclusion that this version of modeling would be better than an observed to expected.

You know, I see what you're saying about the small sample sizes but using the predicted mortality or the predicted readmissions as the numerator is not without its own negative consequences.

Susannah Bernheim: Right. So I mean, I guess maybe if there are specific things that you'd like to hear more about, we can bring that back to the steering committee and, again, this is an approach that's now being used across many of our measures and with the support of many statistical consultants. And it is thought to be the most fair way to assess hospitals and – in a public recording setting that has been a preeminent concern but I'm happy to, you know, work with my statistical team to give some more specific responses if that's necessary.

Risha Gidwani: OK. I think really the overarching question that I have is just for that numerator. Using a predicted value, using – rather than an observed value is bringing in some uncertainty into the estimate because rather than measuring, you're projecting. And so if your team would be able to give their rationale for, you know, why that's still appropriate, that would help me in understanding these to this measure.

Susannah Bernheim: OK. I mean, I will say that in many ways, what we're trying to do is stabilize the estimate right because we're accounting for the fact of the uncertainty associated with small volume and making sure that that estimate is stabilize, but yes. Let me put – let me – let me make sure that we have that in person when we are there in a couple of weeks.

(Lee Anne Hines): Hi. This is (Lee Anne Hines) from CMS. You know, I do understand these questions you raised and as a matter of fact, that every time we bring our outcome measures to NQF endorsement, we run into the same questions.

And as you know that, our several outcome measures using that hierarchical modeling were endorse by NQF for. And I want to back to your questions. You know, just because we got so many questions and got feedback every time we come to NQF, so that's why we consulted five statistical societies and ask them, "Hey, do you have any consensus in your field about statistical model for public reporting or for value – for comparing or profiling hospital?"

So and I also let them know that it seems like every time that we have come to NQF and different statisticians or different researchers, it's the same questions. So we asked them whether they can review and come up some recommendation and they actually recommended the hierarchical modeling and then we provided the URL for that paper.

And after that and I think NQF also convene a meeting and invited the Statistical Society president to discuss about this and I don't know whether NQF members there, do you want to clarify it more or what do you want to – how do you want to deal with this issue whether you

want Yale to have one hour presentation or what? And I just want to know what is the – how do we approach this?

Risha Gidwani: So let me be clear. My concern is not with the hierarchical generalized linear models. I fully support their use. They're the appropriate way to take into account the multi-levels needed for the modeling of the patients in the hospital. I have no issues with that.

My concern is about the numerator being a predicted model rather than observed data. The denominator is expected – is derived from an expected model base off from the hierarchical generalized linear model, totally fine, totally acceptable, I support its use. So the numerator is the issue.

And the issue is not the specific type of model used. The issue is actually for the dichotomy that's been using a model versus using actual data. And that's where my question lies.

(Lee Anne Hines): So you'd expect hierarchical model mean to numerator and denominator. And then I'm not statistician. I thought what we got from Yale and from the white paper from the statistical society is a set, like, that whole rate is based on hierarchical modeling because they have the hospital specific effect add into the numerator.

So you feel separate. There's two issues, then I am lost and I think I need to then, not to answer that question, what a hierarchical modeling can be treated in two sections.

Suzanne Theberge: So, the numerator is based off of one hierarchical and generalized linear model, the denominator is based off of another hierarchical generalized model. And I'm wondering why the numerator uses the model at all. That's the question.

And I did actually – Lisa Iezzoni actually did forward me a risk adjusted researcher from Harvard. She did forward me this white paper this morning called Statistical Issues in Assessing Hospital Performance which was commissioned by the Committee of Presidents of Statistical Societies reported to – submitted to CMS in November 2011.

And in there on page two, they do talk about this issue of observers expected to answer how we really deal with this generalized linear models and the use of the hierarchical modeling rather than the issue of whether you want to use actual data versus the – any sort of model at all for the numerator. All right. Is that clear or should I provide additional clarification?

Female: I just don't – well, I'm just – to provide my – from CMS perspective. I am pretty sure this is, you know, sort of – we're going to have a group of people discuss this and I welcome other members come in. And if you think that you need anything from us or Yale to talk about it from a separate hour or so. You know, it's fine with us but I – you know, I just want to hear what the group says.

Female: So the AAN has a couple – five problems with the model as well and you've spoken very highly towards the statistical modeling and the rigor. And that's great but you're missing a lot of data on the clinical quality side. So, we had a technical panel give advice to CMS against the …

Reva Winkler: Excuse me.

Female: Yes?

Reva Winkler: Hi. Who's speaking?

Female: This is – yes, this is (Sarah Town) with the AAN and we're part of the AMA PCPI and we're the AAN and a lot of the (inaudible) measures. Yes. So…

Reva Winkler: Right. We'd ask you if you could please just at this point let the work group members and the measure developers discuss. The work group really needs to get through the work we're asking.

We'll give you some opportunity for comments and, you know, at the end of the call and then throughout the rest of the process. But right now, we really need to have the work group members discuss this. Thank you.

So, Risha, what is it that you, you know, would help you along here?

Risha Gidwani: The question that would be necessary for me to evaluate this measure is, what is the advantage of using a – any model to – in the numerator rather than using actual data in the numerator?

Reva Winkler: OK. Susannah?

Susannah Bernheim: Yes. So, I think that I'm going to end up saying a lot of the things that I've said already which is that it stabilizes the estimates. It allows us to account for a small sample size as well. If you were a hospital, you would much rather that we use a prior in the numerator rather than estimating you purely on an observe.

But I do think it makes it impressive to put together a formal – a more formal response that I can have my assistant look into that to bring it back. And I'm happy to do that.

Reva Winkler: OK. That would great. So I understand what you are saying about the small sample sizes. I think – and this is, you know, the field of statistics, whenever you make a decision to go down any path, it's always going to have its pros and its cons. So why are the advantages of the model outweigh the disadvantages of the facts that any prediction is going to have on certain (inaudible)?

Do the small sample size – do you feel – does your team feel that the issue of small sample size is for some hospitals being mitigated by this approach outweighs the issue of all hospitals in the sample having some noise built into their estimates by use of our predictive model for the numerator? Great. OK.

Suzanne Theberge: Thank you.

Reva Winkler: Yes. Anybody else on the steering committee have any other comments or questions to add about this measure?

Ramon Bautista: This is Ramon Bautista again. I think it's a good measure – a fair measure. You know, I actually think that the risk adjusted that we discussed in the last few measures. I like to probably hear the AANs comments, maybe not right now but down the future what they thought maybe missing from the risk adjusted factors that they'd like to consider here.

Reva Winkler: Right. OK. Anything else from other steering committee members?

Female: No.

Reva Winkler: OK. Then can we move on and just your thoughts on usability and feasibility for this measure?

Risha Gidwani: Well, this is Risha Gidwani and I have the same sort of conclusion of – I stated insufficient just because I'd like some more information on the modeling before I can assess the usability.

Reva Winkler: OK. Anything from anybody else on the committee?

(Carrie Christiansen): Yes. And for usability – this is (Carrie). I didn't (know) – see them talk about what the plan for using that on the public (inaudible). And although, clearly, there's a system that's saying that it's for that. I guess I just didn't see it properly discussed or – and that will be using QRI.

Reva Winkler: Yes. This is Reva. One question I have for Yale and CMS is this measure is applicable only to the Medicare population, correct? I know that for some of your other similar measures, you've tested it in all age populations. Is that in the works for this one or is there a decision not to?

Lisa Iezonni: This is Iezonni, I have to speak to that. Yes. We do plan to do similar testing. Quite honestly, your call came up much sooner than we were expecting it. And so, the testing isn't complete yet.

Reva Winkler: Yes.

Lisa Iezonni: We can either bring it back as part of maintenance or try to complete it during this cycle.

Reva Winkler: OK. All right. I was just curios if that was happening or not.

(Leanne): Yes. This is (Leanne). You were right. Our goal is to get them – be usable for all ages including for Medicare, so.

Reva Winkler: OK. Well, this is a new measure so, you know, there can be evolution as it – as we move forward. Does any of the steering committee members have any other comments on this measure before we move on to the last one?

OK. So, the last measure is on the 30-day hospital – 30-day all cost risk standardized readmissions following acute ischemic stroke hospitalization. And Risha, I guess this one is yours too. It looks like in your comments there are a lot of similarities between this and the

mortality measure, perhaps you might want to focus in on the areas that are specific to this one that might be different or have other questions or concerns.

Risha Gidwani: Certainly. Thanks. OK. So, this is our last measure. It is 30-day all cost readmission as for acute ischemic stroke and it is a hospital level measure using data from inpatients who are 65 years older.

In terms of the admit – exclusions of patients from this measure, patients who pass (Q.A.) in the hospital were transferred to another acute care facility, were discharged against medical advice. And those that do not have at least 30 days of post discharge data (inaudible).

And this is as well a statistical risk adjustment model. It's an outcome measure using billing data. It also is a hierarchal logistical regression model modeling both patient level and hospital level factors.

In terms of the final variables that were used in risk adjustment, they were those based off of age, gender and a variety of clinical characteristics. These variables were chosen based off of empirical analysis, prior literature and clinical judgment.

I guess from my perspective, I had some of the same questions I had for the other model with one exception being that the area under the ROC curve is I believe 0.602.

Well, you know, and that – I'm just wondering if, you know, what the developers feel is a rationale for using this model. So, the ROC curve indicates the ability to discriminate patients that have readmissions for patients who do not have readmission.

A value of 0.5 indicates that there is no discriminative ability that the model would be roughly the same as a coin flip. So, a value of 0.602 is a little on the low end. So, we'd love some feedback on that.

Female: Hi. Yes, it's a good question. It's one that comes up with the readmission measures fairly often. The way that we build these models is that we only risk adjust for factors that are patient factors and that are present at the time of admission because the goal of the risk adjustment is really just to put hospitals on a level playing field for what to come in.

And it turns out for readmission and this is just not true for our model but for many readmission models that are out there. Those patient co-morbid factors are not a strong predictors of readmissions as they are, say, for mortality.

Our thinking about this is that readmission is really more of a signal about hospital quality and about the transitions of care and about the follow up care. And these are things that we aren't putting into the model on purpose because they're the things that we're trying to illuminate as we look at differences across hospitals.

And so, we have found consistently that the readmission measures have lowered discriminative ability and we think that that really reflects on the nature of what we're measuring. And again, if you look at readmission models that use our clinical data as well,

you know, chart obstructive data or claims data. You'll find that overall they all tend to have relatively low (inaudible) statistics.

Female: Is there – I understand the rationale for not wanting to include variables in the model that needs to be improved such as discharge planning or communication. Has there been any thought to bringing in other structural level for (inaudible) such as potentially, you know, patient ZIP code versus hospital ZIP code or, you know, number of, you know, the – sort of the ambulance pattern in that particular area that may improve predictability of the model?

Risha Gidwani: So, you know, I think – and actually as we're having this conversation, it will be interesting to look at as to show that if you, you know, put in things that you don't want to put in, it makes the model look better but it tells you less about an area.

So, if an area say has dysfunctional ambulance patterns that lead to patients having delayed care and higher readmissions. We want those areas to look worse on this measure because we want to highlight those – whatever are the underlying patterns and systems of care that are contributing to the higher readmission rate.

But it would be an interesting experiment. I mean, I know it's a model, for instance, if you put in complications of care, you get more discrimination but you're also losing your ability to identify that as a poor performing hospital because of this complication, right? So, that's sort of the balance.

Female: All right. And that makes sense. I'm just wondering if things that are beyond hospital control since the unit of analysis (visit) facility are the hospitals. Things that are beyond the hospital control like the patient distance from the hospital or the ambulance patterns if that, you know, even though that's an opportunity for improvement that we'd like to improve. If the hospital can actually intervene on that then it might be a reason to include that in the model.

Risha Gidwani: Yes. It's an interesting question. We have not – I don't know that we have data that would let us do the ambulance suggestion but certainly one could think about the distance. And we – and that's not something that we have studied in our group.

Reva Winkler: OK. Risha, any other issues you'd like to raise on the scientific acceptability of this measure?

Risha Gidwani: Let's see. Well, I think no additional questions from (compliance).

Reva Winkler: Right. How about from other members of the committee?

There's still some questions to be answered but do you feel you have the information you're going to need to have a discussion with the entire committee and determine whether the measure needs a criteria or not?

Ramon Bautista: This is Ramon Bautista. I just think that AMA have some other measure we'd like to add as well. So, I'd like to want to hear the (AAN) as well later on.

Reva Winkler: Yes, we will. We're almost there.

Ramon Bautista: Yes. That's it. That's all.

Reva Winkler: OK. All right. So, anything on usability or feasibility for this measure?

Risha Gidwani: No.

Reva Winkler: Is there anything that's particularly unique or different about the readmission measure that should be highlighted compared to the mortality measure?

Ramon Bautista: Readmission may also reflect, (you know), outside the hospital maybe even more I suspect, you know, or maybe more – if maybe more something to many different things outside of direct hospital control.

Female: One question I have was about the elective readmissions. So, it – let's say if – a patient is, you know, came in for a stroke and then 20 days later is, you know, coming in for some sort of elective procedure. Would that – that would then be counted as a readmission. Is there any thought to actually excluding elective readmissions from the denominator?

Susannah Bernheim: Hi. This is Susannah Bernheim from the Yale team. Yes, we actually – we've put a fair amount of thought into that and this measure has identified a sort of numerous, potentially scheduled follow on procedures. The most obvious being a carotid endarterectomy.

So, the way we structure it is that if a patient comes back within the 30 days for a carotid endarterectomy and they do not appear to be coming for a new acute event which is based on what their principal discharge diagnosis is. Then that is considered a planned readmission – a scheduled readmission and it is not counted as a readmission in this measure.

And there's a handful of other similar procedures that are technical and kind of experts felt might be scheduled within 30 days of an ischemic stroke hospitalization. And those therefore will not be considered readmission.

Female: OK. And I'm not a clinician so I'll defer to the clinicians on the panel and then the group. But would there – if there would be a reason that a patient had a stroke and then maybe 29 days later came in and had a – for a scheduled hysterectomy. Would any surgeon actually do that – a hysterectomy or would they say, "No. You just had a stroke?" or putting your elective surgery off?

Ramon Bautista: They probably depend on the reason for the hysterectomy, I would imagine that if for a pretty) big stroke, probably not. But maybe for a TIA, if it really had to be – had the surgery at that point probably go ahead and do it then. So it would depend on the clinical scenario.

Female: OK. So then we might be including some of those – for the minor strokes maybe included some elective readmissions that are not related to the stroke but they would still be included in the denominator and therefore the numerator.

Reva Winkler: OK. Any last thoughts because we are reaching the end of our call? Anything else from the steering committee on any of these measures?

OK. Then we do want to give an opportunity for our audience members to offer their comments.

Operator, do we have any folks listening in the audience so we can open up for comment from the audience, from the other members of the public?

Operator: At this time, if you have a question, press star one. We'll pause for just a moment for the Q&A roster.

At this time, there are no questions.

Reva Winkler: OK. Anybody else who's on the call who wanted to offer any comments? I think for folks from – (Sarah) from AAN, I think this would be a good time for you to offer your comments.

(Sarah): OK. Thank you and I think (Dr. Katherine) is on as well. But I'll just quickly summarize and then she can finish the clinical part of it.

So, when you asked about the difference between readmission and mortality model, basically, the co-morbidities affecting readmissions are not adjusted for, neither the social or economic variables either.

And then one important thing to think about in the readmission model is that it doesn't account for the most complicated cases that risk for recurrent stroke. So, there's a couple of big issues and in with the readmission issue, the transfers of high risk patients from community to tertiary. That will count against the tertiary hospital.

So, there's some methodology issues on clinical quality. The administrative data you have is great. It's just that it's not adjusting for the quality of care from the physician or clinical quality standpoint versus even the statistical modeling.

So, briefly on the mortality measures, some of the problems with the mortality are that, again, the most – the best predictor is not included in the mode, NIHSS – and you include all kinds of co-morbidities with the exception of NIHSS. And, you know, initial stroke severity is the most important predictor.

In addition, the models that have been used for MI for example do work well but this is stroke. And stroke is a very different disease and it needs to adjust for important factors like stroke severity.

In addition, the model adjust for pre-existing conditions but not adjust for those that happen in the hospital or acute setting or hospital complications. And that would actually be another problem with the readmission model because it's showing in the literature that high readmission is often due to hospital complications.

And then another issue is that the denominator for the transfers – the definition of an acute – the transfer issues vary across critical excess hospitals for example. And some claim that in a transfer issue that it's not really the same acute hospital care.

The variety of ER rooms differ and so the adjustment for the transfer type is not adequate.

And then another real key element in all this is that patient preference isn't adjusted for either. And what that means is that people sometimes choose to die or comfort care. And so – and a lot of stroke is end of life, not preventable or unexpected. They're very sick patients.

So, those are some very brief things that I wish the steering committee would consider. And now I'll defer to (Dr. Katherine). (Dr. Katherine), is there anything that we missed here?

Female: No, (Sarah) that was a great overview. And I just wanted to reiterate the concerns of the AAN with regards to the mortality and the readmission measure. And I know (Sarah) has actually indicated several different issues that have been raised by the neurologist who are scared about this measure and the potential for inadequate risk adjustment and wrong characterization of hospital rankings and quality of care.

The two biggest being severity of illness, severity of neurological impairment with day and age and also this patient preference issue that with neurologists and physicians talk to patients about their preference for end of life care. And they actually abide by patients preferences that they could be "dinged" providing care to patients that are in their best interest and actually follow the wishes of the patients.

The majority of patients who die after a stroke are DNR prior to that death. There are a lot of patients – the majority of patients who have a stroke are the elderly. And so, their stroke causes as you know huge disability and often the preference of the patients and their family is to die peacefully and not have their life prolonged.

Suzanne Theberge: All right. Thank you. Are there any further public comments before we wrap up the call?

All right. Well, thanks for your time, everybody.

Ramon Bautista: This is Dr. Bautista. Can I have those comments made written for a meeting later in the month?

Suzanne Theberge: Yes. We are …

Female: Yes. We actually have a letter. Can I submit the letter that we submitted to CMS in August 2011 or 2010 possibly? And it documents all these issues and if we could have that given to the steering committee, we would greatly appreciate that.

Suzanne Theberge: Yes. This is Suzanne. Please send me that letter and I will distribute it to the steering committee. We'll also be sharing the transcript of this call as soon as we have it. That should take a couple of days but we'll let you know when that's posted on SharePoint and our public website.

Next steps for the committee is just to review the remaining measures in the projects. There are two more work group calls, tomorrow and Thursday. You're welcome to call and listen in to either of those if you're interested. And following those calls, we'll be sharing my preliminary results for all the evaluation surveys and then we'll be discussing them, all the measures in person in the meeting in June. You don't have to fill out the work group surveys for any other measures but please do review and think about the other measures that we'll be looking at in a couple of weeks.

And we'll also be reopening the measures as discussed on this call and – for the developers to make changes. And once we have those resubmitted by the developers, I will post them on SharePoint and let you all know when they're available for your review.

If you have any questions at any time, just give us a call or an email. And thanks everyone for your time today. We appreciate it.

Risha Gidwani: Sorry, one last question, this is Risha Gidwani. When might we be able to expect some of the more feedback on the statistical specifications?

Suzanne Theberge: From the developers you mean?

Risha Gidwani: Yes. If we could get that in – maybe by next Monday, that would be really useful.

Suzanne Theberge: So, CMS, Yale, if it's at all possible, that would be lovely. As soon as you get it to us, we will get it out to the steering committee, so.

Yes. They do have a number of measures – the committee does have a number of measures to review. So, the sooner that you can get us that information, the better so that folks have time to look at everything.

Susannah Bernheim: And this is Susannah Bernheim from the Yale team. We can try for early next week. Part of our team is out of town later this week so maybe it may not be (money) Monday. And once that that forum, the AAN raised a lot of issues that we've done a lot of thinking about. What's the best forum for us to help provide some input to the steering committee on some of these issues?

Suzanne Theberge: Well, you can submit responses in writing to us. I can share the letter with you as well if you haven't – if you don't already have it. And you'll also want to be prepared to discuss it at the committee meeting. I'm sure it will come up. And so, you'll want to be able to address it at the committee.

If you give us a written response, we'll share that before the meeting. But it will definitely be discussed.

Susannah Bernheim: OK. Thank you.

Suzanne Theberge: Any other questions?

All right. Well, thanks everybody for your time. And we look forward to seeing you in person in a few weeks.

Female: OK. Thank you.

Risha Gidwani: Thank you.

Suzanne Theberge: Thank you.

Female: Thank you.

Operator: This does conclude today's conference call. You may now disconnect.

END