NATIONAL QUALITY FORUM

Moderator: Karen Johnson
September 10, 2012
3:00 p.m. ET

Karen Johnson: Hello, this is NQF and this is our first workgroup call for Phase 2 of our Neurology project. So, I want to welcome you guys again to our CDC process. I hope you had a nice summer so far over the summer winds, but here we are on September again and evaluating more measures.

So, first stop, I just want to say thank you, again, for participating in this process with us and for doing the pre-work of looking at these measures and taking a deep dive on the epilepsy measures and carotid stenosis measure.

So before I – before we delve into the measures, I did want – oh, Suzanne is telling me we need to introduce. So, let me hand it over to Suzanne and you introduce who needs to (be introduced).

Suzanne Theberge: OK, if we could just do a quick roll call for the committee and the developers. So, we have on the transcript who's in the call. So, as Karen mentioned the NQF team is here, and for the steering committee, Jocelyn Bautista?

Jocelyn Bautista: Yes, I'm here.

Suzanne Theberge: Great! Ramon Bautista?

Ramon Bautista: Yes.

Suzanne Theberge: Great! David Hackney?

All right. Raj Sheth?

Raj Sheth: Yes, I'm here.

Suzanne Theberge: Great! Thank you.

Bill Barsan?

William Barsan: I'm here.

Suzanne Theberge: Fred Tolin?

Fredrik Tolin: I'm here.

Suzanne Theberge: Great! And then for the measure developers on the line, who's here from AAN?

Female: We've got Gina Gjorvad and Rebecca Swain-Eng AAN (inaudible).

Suzanne Theberge: Thank you. And who's here from AMA PCPI

Female: Diedra Joseph, Amaris Crawford, Kimberly Smuk and we also have ACR staff and Dr. David Seidenwurm will be calling in as well.

Suzanne Theberge: OK, and who are the ACR folks on the line?

Female: Judy Burleson and (Collin Talent).

David Seidenwurm: David Seidenwurm.

Suzanne Theberge: Great! Thanks everybody for joining us today, and with that, I'll just turn it back over to Karen.

Karen Johnson: Thanks, Suzanne and thanks for keeping me on track on the roll call which I always seem to forget because I really want to delve into our criteria again. As you know, we are looking in Phase 2 at measures and they will be measures of epilepsy, Parkinson's disease, dementia – with dementia is going to be diagnosis and symptom assessments and then dementia assessment itself and counseling. So, a lot of different dementia measures and a few of the others.

However, unlike the measures that we looked in phase 1 of the project, only four of the measures in phase 2 had been tested for reliability and validity, and I guess you really notice that as you were doing your preliminary evaluation. So, I believe that Suzanne send out a little e-mail just to orient you on why we're looking at those kind of measures.

Generally, NQF would not be thinking and looking at measures that have not been tested for reliability or validity unless they meet three criteria. The first one is that they are not complex measures. So, in other words, they are not outcome measures or risk-adjustment measures that sort of thing. They would need to know to fill a (gap) area. So, in other words if there are no other measures there are in the NQF portfolio that have been endorsed already on the topic area and that is the case for a Parkinson's, epilepsy and dementia.

And then finally, there has to be some sort of a legislative mandate for their use, and for these measures in phase 2, these are measures that are included in the 2012 PQRS system or program. So, you will be looking at these measures in evaluating all of the criteria except for the – you want me looking at reliability and validity testing for those measures, and you will decide if they should receive time-limited endorsement. If they do, then the developers have agreed to do that reliability and validity testing within 12 months of the endorsement date.

So, there are – just to remind you, to bring you back from our early summer in-person meeting, there are four major criteria that we're having you use to evaluate measures, importance to measure and report, scientific acceptability, usability and feasibility and, of course, the first two are what we have must pass criteria.

And I want to give you just a little bit of – to some of the guidance for these measure – for these criteria because the measures that you have in front of you for phase 2, for the most part are really, they feel different than those measures – the (stroke) measures that you looked at in phase 1.

So, when you think about criteria number 1, importance to measure and report, for – just recall that what we're trying to get to with these criteria is we're trying to get to these things that are most likely to drive improvement in health care quality. So, there are things that really are worthy for donating resources to – for collecting data and for doing things like publicly reporting results.

So, this is sometimes quite different than things that are important to doing practice. So, in other words, there are many, many things that are very important to do but with the importance to measure and report criteria, you are asking, you know, does this thing rise to the level being a national consensus standard.

I also want to remind you just in general that there is a hierarchy of a preference for measures and we went over this back (this) day in your tutorial but just recall that we would prefer, if possible, to see outcome measures that have a really strong evidence-based link, but then after than, you know, outcome measures would maybe – not maybe strong of evidence but still evident and I'm hearing some static on the line.

Is that – can we do anything about that Suzanne?

Suzanne Theberge: If you're not speaking, it's best to put your phone on mute if possible.

Karen Johnson: OK. We, of course, do want and need and endorse many process or structural measures but again, just as a reminder, we would like to have those that are most (crossly) linked via evidence to desired outcomes, and so when your thinking about the evidence sub-criterion, we asked you for process measures which everything is based to our process measures. So, we will ask you to rate the evidence on it's quantity, quality and consistency of the body of evidence.

We prefer that measure developers use evidence that has already been assembled and reviewed and if possible, graded by others and if they don't use either the grade system of the U.S. Preventative Service Task Force system, we ask them to describe the system that was used to grade the evidence.

Just another reminder the expert opinion is not considered evidence, so that is something to keep in mind as you think about maybe potentially some of the grading systems that sometimes we do use expert opinions as underlying some of the guidelines for example.

For quantity, when you're rating quantity, remember you're looking at total number of studies not articles or papers. When you're thinking about quality of evidence, for the most part, what you're thinking about there is related to study design including whether or not they are randomized control trials and the level of (flaws), if you will, how good the study is and you're also thinking about the directness or indirectness to discuss specific measure of interest. And then again, when you're thinking of consistency, you want to consider for magnitude of effects found in the various studies, as well as the direction and decide on how consistent this literature is on particular measures.

For scientific acceptability of measure properties, again, that is a must pass criterion. There are two main sub-criterions under that reliability and validity. Again, both of those are also must pass sub-criterion. And as I only said, for the most part, the measure testing will not be looked at in, you know, phase 2 except for four measures, but for those four measures, you will again, be thinking about reliability and validity testing and you will recall that we allow for testing at either the data element for the measures four level or both. We do accept phase validity of a measure, but we do consider phase validity to be the weakest demonstration of validity.

Regardless of whether there is testing or not on the measure specifications will have to be looked at. So, all of the measures that do not have testing results will still have specifications, so you still have consider whether the specifications are precise and whether if they are consistent with the evidence that is presented to you, and you'll also want to consider things like exceptions to the measures and that sort of thing when you're thinking about validity.

Again, remember that the rating of really – reliability (inaudible) really relate to some extent or to a large extent and how the testing was done not so much the results of the testing but how much testing.

And third and fourth criteria are usability and feasibility and those are what we consider – we do not consider that it is a must pass criteria but they are extremely important. For usability, we're looking to find the measures that are meaningful, understandable and useful for a public reporting and quality improvement for both of those, and then for feasibility, we are looking at basically how easy is the measure to implement.

We're looking at data burden and whether or not data are readily and easily available, and I guess just one criteria or one little caveat here, feasibility is very important, but it is a less important criteria if the evidence is not there. So, in other words, if it matters what – how feasible something is if it's not an evidence-based measure.

And then finally, before we actually delve into our first measure, I did want to give you just a little bit of the results of a CSAC guidance. So, CSAC as you may recall are Consensus Standards Approval Committee. They are the next level after the steering committee looks at measures then the measures go on to the CSAC and about a year – little over a year ago, the CSAC provided some guidance on quality performance measure construction.

And I believe we provided that document to you way back at the beginning of the project but one of the things to – that the CSAC said in that guidance document was they will really hoping I guess that the measures that NQF would endorse would not be measures that are met primarily through documentation without an evaluation of the quality of that activity. So, in other words, this so called, "check box measures" are less of interest and probably less valuable in terms of impacting quality of care.

So, maybe, I'll just stop there, and we can leave this if that's we need to. So, we had asked Jocelyn to start us off with measure (1553). So, Jocelyn, why don't we start and let's kind of go slowly and kind of do the way we did it in the – in the in-person meeting. So, maybe give us a really brief introduction of the measure itself, you know, what it does and then we'll start and discuss impact and go from there.

Jocelyn Bautista: OK. All right. So, my measure is NQF 1953 seizure type and current seizure frequency. This is a new submission from the American Academy of Neurology, and this measure is meant to capture the proportion of epilepsy patients who were seen in outpatient clinic and for whom seizure type and current seizure frequency are documented in the medical record.

So, the numerator would be the patient with the seizure type specified and the current seizure frequency for each seizure type documented and the denominator would be all visits for patient with the diagnosis of epilepsy. There is one equal group of exclusions they (lumped) under documented medical or patient reason for not recording seizure type and seizure frequency such as the patient or caregiver unable to provide the information. This is a process measure.

The submission list administrative claims of the data source for this measure and apparently there is a new CPT category. The submission it says in development – I don't know if that's still current. Sorry, I don't know exactly how that works if there is some code that you need to abstract from the medical record whether this was performed and then translates that into a CPT code. But the level of analysis is at the clinician level.

So, unless there any specific comments about that general overview, I could go to the importance to measure and report. OK, so under importance to measure and report, our first category is impact, and of the four individuals who voted, three rated this as impact, one as moderate impact. There was one comment in the evaluation. Someone asked, is this really an issue. I'm not sure exactly what he's meant by that. I rated this as high. I do think epilepsy affects a large number of individuals can have significant morbidity and Institute of Medicine issued a report earlier this year highlight the need for a national quality measurement strategy for epilepsy.

Fredrik Tolin: This is Fred, Jocelyn, and I was the one who wrote – my comment was this – is this really an issue? And the reason I wrote that I was struck by the fact that this measure is really something that I would think is a third year medical student information. If somebody has epilepsy, when you see the patient making note about any recent seizures seems like the first question that would be asked, and if it's not, then I'm really misunderstanding what this is going after.

Jocelyn Bautista: Right. So, that gets one (1B) the performance gap and so the submission refer to a chart refer of 300 some individuals with epilepsy that showed a compliance rate of about 45 percent for document of seizure type and 28 percent for seizure frequency. So, at least for that particular study, we're not doing a very good job with this. I should have mentioned when I introduced this measure that this measure is meant for the primary care physician primarily not necessarily for the epileptologist.

Karen Johnson: Jocelyn, this is Karen. The – in the submission, I guess one of the things that I would have expected to see would be just a, you know, a fairly quick treatments of impact telling me how prevalent epilepsy is. Was that in there or?

Jocelyn Bautista: Yes, I think so. I have go back and look, again, but I thought that was in there – early treatment cause in epilepsy or maybe (ideal) with this day-to-day (inaudible) for that so much. (Inaudible).

Raj Sheth: Hi. Jocelyn, if I can make a comment. This is Raj.

Jocelyn Bautista: Yes.

Raj Sheath: You know, I think Fred's point and it's a valid point but the problem is – the fact to the matter is that it isn't practically done and whether you have a focal or generalized seizure, we don't know is like whether you are at the North Pole or the South Pole in epilepsy.

For instance, if you have generalized epilepsy, many of those patients do not need MRI scans, do not need imaging studies and require a specific type of treatment. If you have a focal epilepsy, you do need a scan in many patients and you do need the medications that are not likely to make the seizures worse.

So, I understand, it's a – I mean, a fully (read) with you that this epilepsy 101, but the fact to the matter is, it isn't being done and it is certainly increasing cost tremendously. Many patients get put on the wrong medicines and actually the seizures become worse and all of this really goes back to the root structure that the step 1 hasn't been done correctly. And I can understand Jocelyn's point because it is very frustrating when you see this consistently happening all the time.

Male: That you brought – that's – I love what you just said and that really flushes out kind of the skepticism I have about this, but – and in addition, well (inaudible) I'm skeptical about it but I – I'm much less skeptical with looking at it from that perspective.

Male: OK.

Karen Johnson: This is Karen. And maybe Jocelyn, when you get to the evidence sub-criteria, I really want to sit and park for a little while on the structure process outcome relationship, and I think that might be a part – difficulties of submission is that is not really sold out for the non-epilepsy folks on the panel.

Jocelyn Bautista: Yes, so certainly. So, in terms of impact and performance gap, I think I don't know that I have any more to say. I think they do demonstrate impact and a performance gap and four out of four of us – well, if you come both moderate and high four out of four of us agreed on that.

So, moving to evidence then, in terms of quantity, I had difficulty determining the actual total number of studies. They do say that there are two guideline recommendations and several retrospective studies but I don't know that they actually come up with a total number of studies but all four of us rated the quantity is moderate. In terms of quality, I – from my reading of it, most of the evidence is retrospective of observational studies which I rated as low in terms of quality. We were split there. So, of the four in the group that voted to rated quality as moderate, two rated quality is low.

And then I had a lot of difficulty judging consistency but there did seem to be a general consistency that the benefits outweighed any potential risks. But the issue in terms of whether documenting seizure type and seizure frequency whether that (inane) of itself leads to better outcomes. I don't think there is any data there. What they show that – the observational study showed that there is a link between the type of seizure and the frequency of seizure and a patient ultimate outcome. But I don't think there are any studies that show that documenting it leads to better outcome.

Ramon Bautista: Jocelyn, this is Ramon. The part of the reason here is that the issue of, you know, epilepsy (classification) is actually has been (inaudible) for such a long time now. I (presume) – I assume that anybody comes in with seizure need to have – need to have the right documentation on board as far as seizure count and that type of seizure you have. It seems like asking a strong patient what kind of stroke you have.

So, the data will not really be there in the same (inaudible) that data for let's say, you know, MRS for example will be there. But by – it's part of basic practice though to know this kind of information if you are to treat somebody correctly.

Karen Johnson: So, this is Karen. Let's stop there for a minute, and let's look at section (1c1.1) that where we ask the developer to actually tell us what the structure process outcome relationship is because the kind of thing that we're looking for there is to say, you know, what is the measure focus and how do you envision that – (expecting) an outcome and what outcome would you expect.

And I think, Rebecca you may want to speak to this, but maybe they didn't understand what we were asking for here, but I think it would be really helpful for those of us who are not clinicians at all and certainly not epilepsy docs to understand and maybe it's causes as one of you said earlier today. But what is that link – how – what's the link between documenting type and desired (health) outcome.

Rebecca Swain-Eng: Sorry. I (still need) to get my phone off of mute here. This Rebecca Swain-Eng with American Academy of Neurology, and this measure was put together by large workgroup of epileptologist and also representation from not only the academy but other groups but more of getting to your question of what is the link from the actual question of

documentation to – documentation of seizure type or seizure frequency to the improved outcome.

I don't have any specific studies in front of me. I think what you would call this is more of an indirect link but as Dr. Ramon and other physicians have said on this call today that it's very important and it's the best practice.

This is something that we are putting into place and various different registries and various different outlets by different groups for performance improvement practice for (certification) where we are getting some data showing that there is an improved outcome by – just to complete that documenting that they are recording with seizure type or seizure frequency in their medical record.

So, those – the information and the data that we do have today is very limited but coming in so far as it has been very positive, and at the moment it is – I guess what you would call more distal or indirect.

Karen Johnson: So, can you give me any kind of trajectories and let me put out a hypothetical one and you tell me if that's – is it getting there. You document the type that leads to better treatment decisions which leads to better outcomes, you know, lower morbidity, lower mortality something like that. Is that what you're envisioning that's the kind of links that you're saying?

Rebecca Swain-Eng: I think the most important link is that just by documenting probably it's– excuse me, documenting seizure frequency and seizure type, you are able to better know what type of treatments to offer to the patient. That's the most important thing that comes with this.

As one of the earlier steering committee members has said, it's very important that, you know, these two simple things before you can appropriately treat the patient know what type of test that you would do or would not the necessary to do which could then lead to decreased cost not doing unnecessary treatments or procedures that could eventually lead to better patient outcomes, better seizure management, better quality of life.

Karen Johnson: OK. But did I also hear that you say that you did have some preliminary evidence that documenting has led to some improved outcomes?

Rebecca Swain-Eng: (Inaudible) preliminary, we have this epilepsy measures in our American Academy of Neurology performance in practice. It's an online module. It's called Neuro PI where we put the measures into that program whereby physicians can report whether or not they are using the measures and how well they're doing and then give us feedback on how well they think (inaudible) in the (inaudible) of care. It's not very objective data right now. It's still subjective, but we know that the quality of measures are being used by different health care systems.

For example, there is a health care system in Iowa that's adapted them. We're now in California and so on and we're working with them to try and get some preliminary data back from them as to how well this is improving their patient care.

Karen Johnson: I think it would be really important for us to see any data that you have because, I mean, really there is no data in this submission.

Rebecca Swain-Eng: Sure. I can check and see what we can provide with you all.

Karen Johnson: And this Karen again, and I'm interjecting a little bit more on this call than I did in the phase 1 call because, again, the evidence-based for some of these measures is more limited I think than some of those more established (stroke) measures.

So, the really nice thing about section 1c.1 that process outcome link is that really once you can identify what you think that link is then that basically tells you the kind of evidence that you should be (inaudible) seeing in the evidence section. So, again, you know, where is the or what the evidence that this documenting leads to deciding type of treatment which leads to lower cost or better management that sort of thing. So, again, that – it (tells) more (inaudible) roadmap for what you should expect to see further on in the evidence section.

Ramon Bautista: This is Ramon Bautista. The – if you look at the measure, you actually look at two different things seizure type and seizure frequency. Now, I would argue that actually seizure frequency is actually not a process but actually the outcome – an outcome measure because…

Female: Now, but they're…

Ramon Bautista: … the very – the very nature of managing epilepsy is (inaudible) try to decrease their seizure frequency.

Female: But Ramon it's – the measure is the documentation of the seizure frequency. It's not the actual seizure frequency.

Ramon Bautista: I understand that. But to try to prove that seizure frequency is the end of all of epilepsy care but some – that's been standard practice for – I mean, that's the way I treat epilepsy anyway.

So, what kind of documentation do you really need to know to prove that seizure frequency should be a standard for anything?

Karen Johnson: I think – this is Karen again from NQF – really what we're trying to do here is to have endorse measures that are really useful for public reporting and we'll really drive quality and is really I guess worth the resources that folks would have to invest to actually do the data collection, to do the public reporting that sort of thing.

So, again, there are many things that are really important to do and that sort of thing, so what we are really looking for are measures that have a strong evidence base that they rise to that level of being a national performance measure.

Male: Right.

Male: There are two different ways you can look at seizure frequency law. You can see something like OK, are you going to be happy with patients coming in and telling us, "I'm having seizure a month," sort of estimate or one seizure a year sort of estimate. Are you going require an actual seizure count from them?

In this particular measure, at least, to what I'm reading here is that to require an actual count from them to be accurate and then I think that's one of the discussions there should be is how precisely do we want the count to be for public consumption because there is no doubt to my mind of seizure frequency is the outcome – is the measure of epilepsy care, at least one of the – at least one of the major measures anyway of epilepsy care.

Raj Sheth: I don't think – this is Raj. I don't think I hear – I totally agree with you Ramon but the bottom line is that Karen is looking at the evidence and, you know, how this has impact care in a very direct kind of fashion, and unfortunately, all these measures and epilepsy is going to be in the same sort of boat if you will. So, I can see both sides here.

Karen Johnson: Is there – we need to go on. I think to the two different guidelines that Rebecca put in for this measure but the – is there anymore discussion about the link between the measure focused in the outcomes and that piece of it before we delve into what they provided for a quantity, quality and consistency?

William Barsan: This is Bill Barsan. I sort of – I see where – I sort of agree where Ramon is coming from. I mean, I think this one in particular compare to the other seizure one, you know, there is actually – there is actually evidence that treating based on the seizure type is important, right?

Male: Yes, you're right.

William Barsan: You know, there's good data on that, and there is good data that suggests I mean, like Ramon said, seizure frequency is an outcome. So, you know, documenting – if you're not documenting what should be an outcome, you know, that seems to be problematic, as you know, what are your – basing your management on if you're not even, you know measuring what should be a reasonable outcome. So, but I agree with the others too in terms of saying, "Well, what's the evidence of documenting these two things leads to better outcome?" And Iran (inaudible) Iraq and hard places to some extent, but it's a tough one – a tough one to evaluate.

Karen Johnson: And this – again, this is Karen – and this, again, is going to be a little bit of difficulty which the phase 2 measures and because the evidence-base is a little bit more limited.

So, what I need to do is just to draw your focus back to the question to just meet the NQF criteria because that's what we are asking you to evaluate issues (inaudible) just try to keep that in mind as you think it through because it's going to – it's going to be difficult perhaps for some of these measures to think about. Maybe they don't need NQF criteria

but, you know, that they are extremely important today and, you know, that's something that's difficult for them to balance I think.

Let's go ahead, Jocelyn and let's look at what the developer told us. They have two different guidelines, the SIGN and (NICE) guidelines. I guess one of my questions in you epilepsy folks would now about this, are these considered great guidelines or, you know, one is from Scotland and I think one is from England. Are they still great that the U.S. doesn't need to create one or I'm just very ignorant on that and curious about it?

Male: I think the primary problem there is that, you know, in the U.K. and Scotland, the (inaudible) and the (inaudible) and the main (plus) and that manages – the patient is actually the primary care (inaudible) and had not been a neurologist. In the U.S., the majority would be neurologist, epileptologist who would be taking care of patients like this and not primary care. But I think they both saw the same goal and that they're trying, you know, to forget the fact that it's a primary care, they both are trying to address the same sort of issue of standardization in some fashion. Jocelyn what's your thought on that.

Jocelyn Bautista: Yes, you know, I completely agree with that. I think, you know, so there the group rated the quantity is moderate, the quality two rate it as a moderate, two as low and consistency one high, and three moderate and all four of us in the end did think the evidence supported the measure.

Karen Johnson: So, maybe those of you who did the ratings – all of you were able to right in moderate for quantity. Can you help explain how you came to that decision? And it's perfectly fine if you're saying that you know from your own experience, certain literature and such that you know that there is a moderate level of evidence out there, but not everybody will have, you know, experience and expertise.

So, what I'm trying to do here is to understand, you know, what's in the submission and if that, you know, did you have to add your own expertise to that – to what's already in the submission? Or did you go simply based on the grade that was…

Jocelyn Bautista: Well I – this is Jocelyn – so, the submission mentioned at least four retrospective studies, and I remember looking at some chart that NQF gave us earlier that I think it's a three or four studies would rated as moderate or something like that. So that's what I use to grade it as moderate, and on the quality, they – the guidelines were level C. So, I took that as low quality.

Karen Johnson: So, does anybody – let's just do quantity real quickly. So, you're going to the four systematic reviews that they talk about. Did anybody else use that figure to help them assign the moderate rating is that what you used?

Ramon Bautista: Yes, this is Ramon. Yes.

Karen Johnson: And the systematic reviews again just what (inaudible) clear about these systematic reviews. Did – were these reviews that actually showed that documentations improved outcomes or, you know, alternatively that, you know, the documentation….

Jocelyn Bautista: No, no. These reviews simply show that seizure type affects essential outcome.

Male: And that seizure type have their own specific management algorithms to them through that way.

Karen Johnson: OK, seizure type effects outcome.

Male: And management.

Karen Johnson: OK, the path then is seizure type effects management which affects outcome.

Male: Correct.

Karen Johnson: OK. So, going on to quality, let's see. Were they both level C that – the both guidelines were level C or one with level B and the other one level C? Actually, I could see a level B for the NICE and then I didn't see the grade for the SIGN?

Jocelyn Bautista: I think SIGN was four. On page 11, 1c.16 – yes, I see you're right. NICE is level B and then SIGN just right underneath that that is level C.

Karen Johnson: Level C. OK, great. And then Rebecca does tell us (below) what the levels mean, so at the bottom on page 11, this is what the levels mean for the NICE. So, if we look at B, directly based on category 2 evidence which is at least one consultant studies without randomization, so that's an observational study for at least one of the (quasi) (inaudible) study, OK or extrapolated from category 1 evidence which is extrapolated from at least one randomized control trial.

So, it's hard to tell if there was – it's hard to tell how many randomized control trials it may have been extrapolated from or…

Female: Yes, there are no randomized control trials.

Karen Johnson: OK, so everything is observational. OK.

Male: I have a question Karen, but if you look at the criteria for quality, I wonder if quality means large, precise estimates of effects.

Karen Johnson: OK.

Male: Now, this tells you I think in way fulfill that both. You know, even they are not RCTs they actually, you know, at least from a practice point, they are an estimate of the effect of what happens after you made correct diagnosis based on seizure classification or seizure count.

Karen Johnson: You're absolutely right. I mean, a really good observational study can be, you know, even better than a not so great randomized control trial. So, it's definitely not that we're saying it has to be an RCT to get a high or anything like that, you know, if they really get observational studies that have, you know, this tight often it's (interval) in that sort of thing which generally comes from large, you know, sample sizes to be able to have more precise estimates.

Male: Part of the problem, you know, is that I mean, I think in this day and age, 21st century, you're not going to find any RCT on seizure type and, you know, management plan. For the part, these things have been mapped out, you know, decades ago from a treatment point of view. At least for most types of seizures. However, you are going to find lots of studies with indirect evidence that these kinds of medication or that kind is preferable for this or that kind of seizure type and that – and the epileptic drugs of some sort within a certain type of population for example are going to cost us significant benefit as far as decreased seizure count.

So, you're not going to find the RCTs that we did in the first phase 1, you know, meeting a couple of months ago because these are actually very old days that we've been doing for many, many years now.

Karen Johnson: OK, great. Well, that's good to know because (inaudible).

Male: Because every single clinical trial we do for example for epilepsy drugs is in many ways a confirmation of these there's a 50 percent seizure induction for example impose with partial epilepsy would take a certain drug or, you know, that I know of epilepsy surgery is, you know, a 70 percent (applications) for people with this kind of an epileptic condition for example.

So, you do have indirect evidence of that sort. Now, is there one big RCT who's going to say, "If you actually – let's have two groups, one folks you know what kind of (seizure) they have, other one you don't know what kind do you have, and let's see that the outcome is five years from now. No, you won't have that kind of study.

Karen Johnson: I think that's really good information to know because, and again, you as an epilepsy doc will know that that other folks on the panel may not know that.

Male: Yes.

Karen Johnson: So, that's good information. I guess we would do that same thing looking at level C for the SIGN. So, the SIGN is a little bit different. So, basically when you're looking at evidence, and I think we've established for this measure that there is not going to be randomized control trials necessarily that are going to be give you the exact thing that we're looking for here with this measure focus, and I think we've also I think understood from what you've told us that there is probably not any studies showing that documenting leads to better outcome or not documenting leads to poor outcomes.

So, we're relying on indirect evidence here, and we still need to be able to talk about quantity, quality and consistency and, you know, we asked the developer to summarize these for us but often you can, you know, sometimes you can also tell based on the level that are assigned if they have been graded sometimes those levels would tell you that.

So, what I'm hearing from you guys is that these levels are – there's enough information here from the level, the use of comfortable that the – there is a moderate level of evidence showing – here's what I'm getting, I'm still getting a little bit lost because I'm not sure what the evidence is showing, but there is a moderate level of it, I guess. Can you guys help me out on this a little bit?

Male: Well, Karen, I think you're – I can clearly understand from a measure's point – from a quantitative of view that this – it's very difficult to show any numbers that support the case and so you could easily (a lot) of you – if you can't actually provide evidence that supports the case then how does that become an NQF realm of practice that should really more of an AAN type of practice where, you know, that's the standard of care. You have to classify seizures and things like that.

But on the other hand, when you look at these measures and you look at something like seizure classification, if you do not put them in the right box, they're going to go down pathways that are definitely going to impact quality of care, and I think the group probably is going to have difficulty with all of these measures in terms of squaring those two very consistent but differing points of view, and I don't know if that sort of help so. Do I – should you correct some of the things I said perhaps give me an idea where you are coming from Karen?

Karen Johnson: You know, I think…

William Barsan: Yes. This is Bill Barsan. So, Ramon, I think one of the other issues is that just because you're assuming that if somebody knows these things that they are going to act appropriately and do the right thing and just document it and doesn't mean what to do with them either.

Male: Right. I mean, if you're an epilepsy neurologist, yes, I would assume that you are going to do that, but if you were looking at these for primary care physicians as well, I'm not certain that just because they document the seizure type that they're going to know – that that the fact they'll mean, they know the proper treatment for that patient.

Male: Yes, you're actually right, (inaudible).

Male: We'll implement the proper treatment.

Male: (Inaudible) under feasibility and usability, we can discuss it in more detail, but you're actually are correct in that point.

Karen Johnson: OK, I'm kind of taken over your job here Jocelyn. I apologize, but why don't we go ahead and move on to scientific acceptability. This is very quickly. Was there anything

there that you wanted to point out or because I know Dr. Sheth needs to leave at 4.15 until I…

Jocelyn Bautista: Will there – I mean, there is no reliability testing or validity testing submitted, and in terms of the measures specification, I don't think even those are clearly outlined. So, again, this whole issue of the CPT code, you know, who determines whether the measure was met, what is the definition of the definition being met. You know, mentioned is it actually documenting a number of clear frequency versus saying, you know, there is a seizure last week or something that – I don't think that level of detail is in the submission.

Female: (Inaudible).

Jocelyn Bautista: But I read it as an insufficient in terms of the reliability and validity and sufficient information.

Rebecca Swain-Eng: This is Rebecca. Can I just make one point of clarification or should have been CPT category II code that would have been in the applications that's not that is completely (inaudible) oversight with not getting in there but this measure does have a CPT category II code which would be just put on extra line on the billing to indicate that you did the measure that is what the (inaudible).

Jocelyn Bautista: Who determines that the physician?

Rebecca Swain-Eng: The AMA – no, the AMA has a group called the performance measurement advisory group. They….

Jocelyn Bautista: No, no. At the time of the clinic visit, who puts that designation that it was done or not?

Rebecca Swain-Eng: The clinician would note in the medical record and then they have to make sure work with the biller that it gets put on the billing form.

Jocelyn Bautista: And what's the definition of current seizure frequency? Is it just any mentioned of seizures?

Rebecca Swain-Eng: It's going to be users but they're currently having – or since their last visit.

Karen Johnson: And I think, just to clarify here, this measure and the CPT II code that this is going to – has nothing to do about whether this is an accurate count or not is the same that I have documented the number or whatever that number is.

Jocelyn Bautista: Yes, so it's really the physician determining for himself whether he met the metric or not.

Karen Johnson: Right. And I think you may recall, there were several of the – (stroke) measures and phase more than that were documented with the CPT II codes.

Female: Yes.

Karen Johnson: You may not remember…

Female: Yes, no – yes, I remember.

Karen Johnson: OK. And that is – Rebecca, that is on your submission form. It's on page 14th – sorry – 13 of your form. Yes, it looks like it's CPT II code 1200F.

OK, how about usability?

Jocelyn?

Jocelyn Bautista: I thought this is similar scenario. I mean, they say that this metric has – is being currently used but they don't actually provide any information about how useful it has been. So, I rated both usability and feasibility as insufficient. There is someone…

Male: Yes, I did that Jocelyn. There's something about – I mean I – this more of a question than anything else is if you use this metric to compare – let's say different practices. Would you be (inaudible) for example if your practice show that you had more seizures than the patient population you saw compared to those had less seizures, you know (inaudible).

Jocelyn Bautista: But then again, if the metric is simply whether you're (inaudible)…

Male: … programs would be penalized for that compared to let's say a more, you know, first or second leveled program.

Jocelyn Bautista: Yes, I think the way the measure is written it's simply whether you document it.

Male: It's not – it's not – it's not the (fugitive) thing and – because that could potentially happen from a payer point of view.

Jocelyn Bautista: Yes, they're not looking at….

Male: …, you know, they are not looking at…

Jocelyn Bautista: … actual frequency.

Male: … (inaudible) with a care for epilepsy patient. They continue to have X number or Y number of seizures for…

Jocelyn Bautista: No, they are not looking at that. They are not looking at actual value…

Male: (Inaudible).

Jocelyn Bautista: … of seizure frequency. They are actually looking – just looking to see whether it was documented – any frequency was documented.

Male: OK.

Jocelyn Bautista: And then someone also – for feasibility, someone asked the question would this apply only to epileptologist, neurologist or all physicians who care for epilepsy patients.

Male: (Inaudible) as well, you know, (inaudible)…

Jocelyn Bautista: I suspect it's all physician.

Male: … whether or not…

Jocelyn Bautista: Yes.

Male: … non-specialist would be able to actually do this. You know, we've been trying to figure out if somebody is having a significantly difficult at best times. So, for those without much experience, they can actually make an accurate seizure count?

Jocelyn Bautista: Right. And the way I read the measure, if anyone who list epilepsy…

Male: Yes.

Jocelyn Bautista: … yes, who bills for epilepsy.

Male: Right. So that will be the (inaudible)…

Jocelyn Bautista: You are expected to do that.

Male: … the accuracy of the measure. I would probably think might be not as precise with those who don't – were not specialist for example.

Jocelyn Bautista: Right. You know, the whole issue of accuracy is not taken into account at all. This is whether – there's any documentation at all, the seizure type.

Male: Right. But it can be misleading though. For example, if a non-specialist would actually undercount seizures and manage, you know, manage care based on the undercounting of seizure. So, it's really one of those garbage and garbage out for those situations.

Jocelyn Bautista: OK.

Male: So, I mean, if you're going to make a measure like this, you must have at least a reasonable amount of confidence that the measure is going to be an accurate reflection of what really is in fact going on.

Karen Johnson: Right. That's why we need the usability and feasibility data to really know.

Male: Yes.

Jocelyn Bautista: So, in the end – so the preliminary assessment of criteria met (suitability) for endorsement of the four us who voted – one voted yes and three voted no and some of the comments, it's clear that the type of epilepsy and frequency is important information to gather on all patients. It's not clear that documentation leads to better outcome.

Someone else writes (inaudible) intention to measure, describe needs to be improved upon, there's a need to address it's usability and feasibility. Someone else writes I'm not convinced it's useful measure open to discussion and then my comment was it's not been – the measures not been tested for reliability, validity, usability or feasibility and I don't – it's clearly an important thing to do clinically.

I don't think that's an issue whether it meets the criteria again, this is Karen has been saying the whole time – whether it meets the level of being nationally endorsed as a performance measure. I'm not sure.

Karen Johnson: OK, you guys have had some great discussion on this measure, and we spent a lot of time on this one measure. It's quite frankly this type of discussion particularly about the evidence is going to be kind of being in play not just for the epilepsy measures but really for the Parkinson's measures and many of the dementia measures as well.

So, you know, get ready to and cast them really good conversations in the in-person meeting, but with that, I want to go out of order here so that we can hear from Dr. Sheth. If you will introduce Raj's measure 1814?

Raj Sheth: Yes, thank you. This is measure 1814, again, proposed by the – sponsored by the AAN and it's a process measure the data source study level and the level of the analysis is really at the clinician level and the – just as some background before we go any further, patients with epilepsy particularly women with epilepsy are being terrified of having children getting pregnant.

We know from the treatment perspective that the medications as a perception amongst patients with – women with epilepsy that in fact that medications may damage the baby or injure the baby or cause malformations, and in fact, we do know with the pregnancy related outcomes which, again, not addressing in this measure that in fact there are medications that raise the risk of serious life-threatening malformations by 10 to 15 fold in some – in the case of some medications. So, that's a background.

What we are addressing here is a totally different practice parameter-based measure that is really related more to some counts of that not medication related but really related to two issues – one is the impact of the epilepsy on pregnancy and the second the impact of the pregnancy on the epilepsy. And I'll take the first one first, if you look at the importance of the measure based on report, everybody agrees that this is a very important clinical issue and clearly a high-impact area and that's sort of thing.

So, just to give some perspective, there about a half a million of women with epilepsy in the U.S., so it's a big issue in terms of the epilepsy world that are in the childbearing age.

This doesn't include prepubescents and the elderly with epilepsy, so just half a million women with epilepsy in the childbearing age that's impacted.

Now, if you look at the main study that that's provided is really the practice parameter, and it's really entitled, "Management Issues for Women Epilepsy – Focus on Pregnancy" as evidence based. The – so let me stop there and ask about impact and if there are any questions with regard to that.

It seems that there is none. I will move on to the second issue which is evidence and here the – there are a total of 200 – I'm sorry – 876 published abstracts in the literature related to women with epilepsy and pregnancy out of which 285 of them actually were in language that could be – that was English and the excluded patients from this 285 patients that were related – that had eclampsia-related things which we know clearly are not related to the epilepsy as such but really more to pregnancy.

And the evidence is all retrospective of, you know, population-based studies that are retrospective. So, these are not randomized controlled trial. None of these are randomized controlled trials. A total of about 25 of the studies that are reported between 2000 – sorry – 1990 – 1985 and 2007, I believe, where 25 articles had met this and nine articles actually met or superseded class 1 evidence.

Karen Johnson: And Raj, can you help me out here, I'm just looking through the measure submission, and I want to make sure I have that highlighted. Can you tall me where in this submission you're seeing that information in.

Raj Sheth: In the submission, this is the data that's behind the submission. So, they have not actually provided this as best with (inaudible).

Karen Johnson: OK, so this is coming from your expertise of epilepsy.

Raj Sheth: This is coming from the source document from the AAN which is a management issue for women with epilepsy.

Karen Johnson: OK.

Raj Sheth: Rebecca may be able to provide that evidence – provide some of that detail. Rebecca?

Rebecca Swain-Eng: Are you talking about the most recent women with epilepsy manuscript?

Raj Sheth: Yes.

Rebecca Swain-Eng: Yes, I can provide that to everyone.

Karen Johnson: So, Rebecca, this – and Raj, this information maybe we need to go back and just make sure that we understand the process outcome link that is being proposed with this measure, counseling does what? It helps women…

Raj Sheth: Yes, that's a – that really gets to the heart of the matter. So, this is really observational data. There are no issues related to counseling. The only issue related to counseling is if a woman who is childbearing asks what is the risk of me becoming pregnant. What's the risk to my seizure frequency and what's the risk to the pregnancy?

We can sort of say that there is moderate level of evidence to suggest that there – that – I'm sorry – there is a – on the – on the – the effect of the pregnancy other than if a woman is smoking that does not appear to be any increased risk to the pregnancy as such or significantly substantially increased risk to the pregnancy as such.

Karen Johnson: I (inaudible)…

Raj Sheth: (Inaudible).

Karen Johnson: … go ahead.

Raj Sheth: Does that address the issue you're asking or?

Karen Johnson: Well, I'm still trying to understand exactly the link. So, I were a woman of childbearing age with epilepsy and I get counseled about this, what do I expect in terms of my outcomes? Will that…

Raj Sheth: Well, I think – this is more a reassurement type of outcome. So, it wouldn't be – if the question really would be the counseling would be, should I get pregnant if I have epilepsy?

Karen Johnson: OK.

Raj Sheth: That is really where the issue is.

Female: No, but isn't also the whole issue of oral contraceptives being less effective in the setting of certain anti-epileptic drugs, isn't that also part of this? I mean, there are a number of practical…

Raj Sheth: Yes.

Female: … counseling points that affect management.

Raj Sheth: Yes, there are.

Male: But – and then also just using folic acid, you know…

Female: Yes.

Male: … even the (inaudible) variations in folic acid. Isn't that an effective counseling as well?

Female: Yes. I think that all falls under; this is a very broad…

Male: Yes, so we now (inaudible).

Female: … category of counseling.

Male: Is OK or not OK? I mean, you might decide not become pregnant for example or you might decide to change your medications for example or you might decide to take high dose of folic acid. I mean the outcomes are not just whether or not the pregnancy is going to pursue or not to be pursued.

Male: Those are very good points.

Karen Johnson: OK. Did you talk about the performance gap? Was there a question about that or that was from the (pure) article as well, right?

Raj Sheth: Yes.

Karen Johnson: OK.

Raj Sheth: Yes. So the side effects of method of looking at the measures – acceptability of the measures properties, one person rated it high, the other one moderate, and none were low and the validity was rated as two as moderate and two as low.

Female: How was anybody able to really rate reliability on this particular measure?

Raj Sheth: That is – there may be some static on the line again. So that's a good question.

Female: I didn't see any data under liability, did you?

Male: Neither did I.

Female: Yes. I agree that this is insufficient.

Male: Yes.

Karen Johnson: Yes. This one is a little tricky again because they – you know, we know that they do not have reliability testing and we will clarify exactly how to think about this in the (in-person) meeting. But you are able, you know, at the very least to look at this text and think about the exclusions.

Female: Well, that's the thing even if the most specifications are very vague. What exactly is meant by council at the bottom? Let's see. There is no operational definition for what they mean by that. Is the counseling had to be documented by a physician? You know, there's – I didn't see any of that detail provided in terms of the measure specifications.

Male: I think that's a ; big issue for me too. I mean I think, you know, does this mean you give the patient some literature in the office? Does this mean you give them a Web site to look

at? Does this mean you actually spend time talking with them or – you know, what does that mean?

Female: Yes, exactly.

Male: I mean any of those might be acceptable actually but…

Female: Yes and seeing it as a CPT code again. Yes, if the physician, you know, decides from self whether he did the appropriate counseling.

Male: Right.

Karen Johnson: So I think what you're getting at here is actually a different point that underlies why NQF is really sometimes would prefer to see outcome measures and – you know, when we can't have process measures that are proximal if possible to outcome measures because if the further out you get just the harder it is to know that what you're doing in the measure focus is actually going to improve quality and can even differentiate quality between providers.

Male: You know, I think that's a difficulty with this one. And this one too, I mean it is so clear that this is important. And it's so clear that it's a really critical item. What's not clear is that this performance measure in anyway is going to indicate well. So I'm not clear what the performance measure is. But even if we – even if you can define, I'm not so sure that it – you know, that that means there is better quality care.

Karen Johnson: OK. Raj, I know you have to leave in a just a very few minutes. Do you want to bring up anything else on either of the usability or feasibility of the (inaudible)?

Raj Sheth: I think the usability should be relatively easy in terms of, you know, using it. It is just basically these are going to be counseling. But I think some of the questions that were raised earlier in terms of what is counseling exactly. It needs to be defined. I think that we definitely need more information from the sponsor on that.

The feasibility is simply – it's a relatively easy thing to perform but the issue is really how do you collect the data. This goes back to that data collection type of issue and what is the (inaudible) for completing the data in a form that is easy (abstractable). So very insufficient evidence was provided and some of the comments actually highlight this issue as well.

Male: My question, Raj, was how would you know that was the yearly medical records that you're looking at? I mean would you count the dates and – you know, I mean would you go to the trouble of doing all that to document this?

Raj Sheth: You're right this is – it's – the feasibility, I think, if you look at the score that was recommended there it gives the – there were two moderate, one low, and one with really not in the range.

Karen Johnson: This is Karen, the developers have specified this measure right now for both claims and for paper records. So obviously, doing it for claims would be pretty easy. So you question is how easy is it to be for paper records?

Raj Sheth: Right.

Karen Johnson: OK. Anything else, Dr. Sheth, that you want to bring up on this? We kind of breeze through the evidence piece there. Do we need to go back and pick up anything on that piece?

Raj Sheth: I think – really, I don't think on that issue itself. I think the sponsor needs to provide more information for the usability and feasibility issues. And then if you look at the criteria for endorsement, it was equally split. And I think it reflects some of these deficiencies that we've just talked about.

Karen Johnson: OK. Great. Any other discussion from any other workgroup members on this measure? OK. We'll take that one off. And Dr. Sheth, thank you so much for rearranging your schedule to be able to participate this afternoon.

Raj Sheth: Thank you, Karen. I appreciate it. Thank you, folks.

Karen Johnson: OK. Since we've got – we're thinking about epilepsy, why don't we go ahead and just do the third epilepsy measures, the 1954 measure. And Ramon, that one is yours.

Ramon Bautista: Thank you very. So the 1953 actually has sister recommendation to what Jocelyn just gave us – gave us a while back 1954. And so I would like – I'm sorry, 19 – 1954. I just want to make sure my numbers are right. And so just to – just to explain this to the non-epileptologist in the group. There is a subtle difference between classifying seizures and classifying epilepsies.

Seizures (differently) represent the, you know, things like the clinical symptoms, for example, a convulsion as a seizure or a petit mal as a type of seizure; whereas, epilepsy refers to those different disease phase that actually can produce seizures. And then current, this actually is something in the (inaudible) right now.

But we currently actually classify epilepsies actually to being either secondary or symptomatic, for example, from a stroke or head trauma or from being cryptogenic where the underlying cause is, you know, is not really well identified in imaging studies and idiopathic which means that there is actually a defined epilepsy syndrome.

So for example, (absence) epilepsy is a type of epilepsy, an idiopathic epilepsy so is Rolandic epilepsy. So the aim of this measure is actually to try to document in your records whether or not – to check whether or not your records actually the kind of epilepsy type or epilepsy syndrome that you might have. So we are going to know that all patients with a diagnosis of epilepsy and try to find out – are they – you know, if the underlying cause or etiology of epilepsy identified and documented.

So in many ways, the information used by Jocelyn to cover her measures is actually used here as well – from an impact point of view – from an impact point of view, you know, three of us believe that there is a high impact, one also believe there is probably a moderate impact. And from a performance gap, one of us believes that there is a high performance gap and two of us believe that there is probably a moderate performance gap, and one said there is probably incomplete evidence.

Many studies actually looking at epilepsy syndrome also look at seizure type sort of like hand in hand. So it is no surprise at the very same studies quoted by Jocelyn could also be used for something like this.

Is there any question so far about impact and performance gap from the group? OK. (Going ahead then).

Jocelyn Bautista: So you know they – Ramon – Ramon, this is Jocelyn.

Ramon Bautista: Yes, go ahead, Jocelyn.

Jocelyn Bautista: So you know, they used the (inaudible) 2011 study for the performance gap.

Ramon Bautista: Right.

Jocelyn Bautista: But that study actually does not include this particular measure for my reading.

Ramon Bautista: I didn't realize that. I thought it was actually both epilepsy types as well. Was it not there?

Jocelyn Bautista: No. It only talked about seizure type and seizure frequency. So I mean, you know, I just wish you could draw it from your own knowledge of the literature. But based on the submission, I don't know that there was sufficient data submitted regarding the performance gap on this particular measure.

Ramon Bautista: Particular measure, OK.

Jocelyn Bautista: But I mean I think we know from the literature this is an underreported thing. It is not documented as well as it should be.

Ramon Bautista: OK.

Karen Johnson: Maybe we can ask Rebecca. Rebecca, did – I think you said on this measure that you put in the numbers for seizure-type documented. Was that a typo or do you have something also you can…

(Rebecca): (Inaudible).

Karen Johnson: OK. So we are going to reopen your measures so you can go in and have information that the workgroup has asked for.

Ramon Bautista: OK. Now as far as evidence is concerned, you know, it is not a health – it's actually a health outcome, not a process. So the quantity of the data, you know, these had moderate ones that low – the qualities, three said moderate, one said low. And again, it's very much the same – almost the same information as we discussed during Jocelyn's presentation a while back.

And as far as scientific acceptability, you know, we – I agree that, you know, reliability and validity testing was not really completed for this particular measure. Two of us vote it may be moderate, two insufficient for both measures there.

Any comments so far on reliability and validity testing for this measure?

OK. For usability, three of us felt it was probably a moderate acceptability and one said insufficient. And for feasibility, two of us felt it may be moderate, one low, and one insufficient.

And again my comment here was whether or not the measure was really meant to be implemented by epileptologist, neurologist, or non-neurologist because I think even among non-epileptologist, there is a bit of confusion on how to actually classify seizure types – epilepsy types, I'm sorry – maybe a problem especially for the non-specialist.

And as a preliminary assessment of the measure for endorsement who said yes and who said no again, like I think – like Raj highlighting the lack of information and documentation for this measure.

Karen Johnson: This is Karen again. You know, it is interesting I find epilepsy very confusing. I kind of understood just a little bit but epilepsy is harder for me to try to get my brain around. So when you're – when you're talking about documenting the etiology of epilepsy, can you just fill in the blank there for the path? So knowing what the etiology is, does that determine the treatment again?

Ramon Bautista: In many ways, it does. You know, but knowing – knowing the etiology by itself actually requires many things. It might require imaging studies. It might require EEG information. It might require knowing the clinical syndrome or the clinical presentation, and to an extent, yes, it does.

For example, talking about like idiopathic epilepsies, for example, generally are not surgically remediable epilepsies. Whereas, for example, some forms of symptomatic focal epilepsies might be. So that is actually one factor that actually influences treatment types. Even the choice of medications, for example, will depend on you epilepsy syndrome.

(Fred): This is Fred. Is there any documentation that supports that? I mean it (teleologically) makes sense if you can identify the etiology than the treatment plan flows from that and you would expect improved outcomes. Is there any day to support that?

Ramon Bautista: Many – many treatments (inaudible) epilepsy. For example, clinical trials is not, for example – you know, studies of surgical outcomes, for example, are you going to probably target the certain epilepsy type – epilepsy syndrome, for example.

So for example – I'll you an example here, let's say the – a condition called mesial temporal sclerosis, for example, is a kind of epilepsy syndrome that is remediable through surgery. And it has been studies that, for example, a certain percent of patients with that condition are going to benefit form surgery.

And as a corollary to that there is – there are actually medications, for example, that seems to work better, for example, for certain epilepsies. For example, a drug called Zarontin is probably better for a condition called absence epilepsy.

So at this day in age, it has pretty much been – I mean (inaudible) generally tailored a certain treatment options seem to be preferred for certain epilepsy syndromes.

(Fred): Clearly, that's not the – I mean that's not what this is measuring. But I think a corollary it is implied in this step. By evaluating this or looking at this measurement, the expectation is that there would be improved outcomes. And going back to one of those statements that was made earlier by (Howie), we really are looking for and prefer measures that provide some sort of link to an outcome.

Ramon Bautista: Right. And from a practical point of view, (you mean), it has happened quite frequently – I'm sure Jocelyn will contest this as well, is that is so often – it is often, for example, that the wrong medication becomes prescribed for a patient because the actual epilepsy syndrome has not been well validated or well described.

Female: (Inaudible).

Ramon Bautista: So for example, you have patient, for example, who mistakenly get drugs that actually can worsen their seizure type – their seizures – their seizure frequency because the medication they are prescribed are actually contraindicated for their epilepsy syndrome. So actually that – those kinds of thinks are well known and written about.

I mean I think what's happening is that when you're – our literature review for this particular – for these measures actually going to reflect all that. If you really look hard and fast, I mean, there are thousands of studies out there that actually, you know, talk about this different issues.

Male: So here's another thing I'm not clear on, Ramon. This – you know, if you got somebody who is documenting that this is idiopathic epilepsy or whatever…

Ramon Bautista: Yes.

Male: This is because a document that doesn't, you know – you're making two assumptions, one that they're right; number two, that if they document it and they are right that it actually influences their treatment. But neither of those do you really know.

Ramon Bautista: You're right. Your first point is actually very well taken though. It is that my concern again, like – as I said, is that it is very difficult for a non-specialist even the general neurologists do need to be correct in trying to get the seizure type – the seizure syndrome down path much less the non-neurologist.

Male: OK.

Ramon Bautista: And you're right. But it is well known though that once you document – once you have the right seizure syndrome – epilepsy syndrome diagnoses, the seizure – the treatment options are well defined and then that (prove) the outcomes have been well described as well.

Karen Johnson: So Ramon, I'm going to have to leave you as the last epilepsy person on the call.

Ramon Bautista: (Inaudible).

Male: And then – and then – yes, the only other thing than about this – well, I mean, you know, there are plenty besides the things I mentioned, but the only other thing is – so is this something that you – I mean that should be recorded with every patient visit or once a year, you know? I'm just not even sure how (inaudible).

Ramon Bautista: You're right. Your right. It's probably reasonable to say that at least the seizure – the epilepsy syndrome should actually be documented. (Inaudible) this every year – every year, not necessarily every time they come see you.

I mean seizure (cuts) you can understand because I (explained) with you it's probably an outcome measure seizure films as what you're trying to do to improve in anyway. But the actual documentation of what the syndrome they have is doesn't need to be there all the time.

Male: This one – this one was really more of a stretch for me than the other two. I mean mainly for the reason I mentioned just in terms of, again, not knowing whether just because they document a seizure type doesn't mean it is correct and then again making the assumption that if it is correct that it is actually deriving appropriate treatment.

Ramon Bautista: That's right.

Male: And I don't know…

Ramon Bautista: And I mean theoretically they are correct seizure – the correct epilepsy diagnosis actually drives a treatment paradigm for the patient. That's – in many ways, it actually should – should do that. I see the intent of this measure here. And again my concern that I share with you is that it might be wrongly used, and number two, it (might) doesn't need to happen every time you see the patient. I mean why are – we derive every – why do we need to derive this same thing every time we (inaudible).

Male: Yes, but I think it would be worthwhile – I mean if there were any evidence to suggest that. In fact, you know, when the proper seizure type is – or when seizure type is listed that actually derives better treatment in terms of, you know, seizure frequency as an outcome or something.

Ramon Bautista: I mean for the developer – I mean both for this measure and the other one that Jocelyn just, you know, discussed a while back. I mean your evidence doesn't have to just always be RCPs or big position thing. But you can actually look at lots of data out there that shows that certain treatment paradigms are more efficacious for certain kind of seizures or epilepsy types or decrease of seizure count.

I mean those things are out there. I mean the (inaudible) does those kinds of articles out there. You don't need – you don't need big position papers for the opinions all the time. And these data, they are – they are out there.

Male: Right. But I don't – but is there a data out there, for example, that says when you look at patients who were diagnosed with this etiology of epilepsy that in fact the treatment that is being prescribed for them is what's recommended.

Ramon Bautista: Oh, yes – oh, yes, definitely. Definitely. In fact, you can also – there is also (inaudible) that shows – I'll give you an example…

Male: I don't mean that – I done mean that you the treatment that should be recommended. But is there evidence that the treatment that should be recommended is actually being given when the etiology is correctly documented? You see what I'm saying?

Ramon Bautista: Yes, I (inaudible).

Male: I don't believe they (inaudible) what should be getting.

Ramon Bautista: I don't (inaudible) if you don't know these things. You hope it is. You hope it is.

Male: Yes.

Ramon Bautista: But it is evident that it is not being given. I don't know.

Male: All right.

Ramon Bautista: I mean these are important measures. You know, at the end of the day, if quality is what you're looking for, these are important measures. But you know – and I guess I believe that when it actually becomes meaningful and useful is a more important thing.

Karen Johnson: This is Karen again. You guys are really, I think, struggling with something that is going to be the theme throughout Phase 2. And again, it really emphasizes I think even as you, Ramon, has mentioned, you know, maybe even knowing, you know certain outcome measures might be a better way of differentiating quality and improving quality, and we'll

probably include that in that (little section) about future measure development when we write our report.

But since we're getting close to the end of our time – boy, an hour and a half has flown by, and we have one more. And so if nobody has any other comments about measure (15)54, let's go ahead to the carotid – what is it – stenosis measurement in carotid imaging studies, and Dr. Hackney, are you on the line?

OK. We had asked Dr. Hackney to be our lead discussant that in fact something has kept him from the call. So I'm going to put this out for possible discussion by either Bill or Fred? Do either of you feel comfortable being the lead discussant for this measure or do you want me to do it?

Male: I don't know – yes, I'm not sure that I – this is a little bit outside of my – I mean I understand something about it, but not quite up my alley.

Karen Johnson: OK.

Male: (Inaudible).

Male: I would have to say to Karen, you probably are more familiar with the data behind this if you're comfortable with that.

Karen Johnson: OK. Why don't I give it a shot?

Male: OK.

Karen Johnson: And you guys can come in.

Male: OK.

Male: All right, sounds great.

Karen Johnson: OK. This is a measure – this one is a different measure. It's obviously not epilepsy but it is also not a brand new measure. This has been endorsed for quite a while. And according to my information, this is still on time-limited status, which is kind of odd because usually time limited this should be a 12-month thing that – so this has been around for a while.

The numerator – well, the denominator is all final reports for carotid imaging studies that were performed and the numerator is that the final carotid imaging reports include direct or indirect reference to measurements of distal internal carotid diameter as the denominator for stenosis measurement and I really don't know what that means. I mean I kind of know what it means, but I had to look up what stenosis means before we started this project.

There are no exclusions to the measure. It is a specified at the level of analysis for clinicians. And data source is specified as administrative claims, electronic clinical data, and registry data. And it is put forward by AMA-PCPI.

And since we have folks from PCPI who has developed the measure, maybe it would be really good for you to give us just a little bit of a background about why the indirect or direct reference of distal internal carotid diameter is important. That might help us kind of start the discussion.

David Seidenwurm: Can I take that? This is David Seidenwurm. I'm the co-chair of the committee that developed the measure.

Karen Johnson: That would be great. Thank you.

David Seidenwurm. Sure. OK. The reason we chose that terminology was at the advice of our methodology consultant who was a vascular surgeon and we wanted to capture the definition of stenosis what we've used in the randomized controlled trials that documented the efficacy of endarterectomy and that was the NASCET study.

And so the NASCET method, as it is called, is – for measuring stenosis is based upon the ratio between the diameter of the narrowest segment and the diameter of the next parallel-sided distal segment of the carotid artery.

And the reason that was chosen by the NASCET and this was documented in some of the methodology studies that we're published at the time that the study was embarked upon is that this is much more reproducible than the other – what the other competing method was at that time which was – is now called the European Method in which one would infer the diameter of the carotid valve as it might have been when the patient was young and healthy. So because that was kind of an influence, the direct measurement was – of the carotid – of the distal carotid was used.

Then with respect to the term indirect measurement, because of some of the methodologies that are used in noninvasive carotid studies are – the method is indirect. One doesn't measure the diameter of the vessel directly. For example, in ultrasound, one infers the ratio of stenosis through velocities and velocity rations and the – and then validates that other against other studies such as angiography or CT angiography where one can make this direct measurement.

Similarly, you have a problem on the other side with certain MRI techniques less often used today but still in use where one simply cannot make the direct measurement of actually the narrow segment what makes the – what makes an indirect physiological method based upon the signaled characteristics.

So that's how we came up with that terminology. I wish we were allowed to say use the NASCET method because then everybody would know what that meant.

Karen Johnson: So…

Male: So – go ahead.

Karen Johnson: I just want to make sure that I understand this. So the point is that you want an actual measurement of the carotid diameter and you're not saying it has to be direct or indirect. They can be either and that would work.

David Seidenwurm: That's right. That's right. It is just as long as you're reference point is the distal carotid and not your intuition as to where did the carotid valve might have been.

Karen Johnson: OK. Any questions from the workgroup numbers about what our experts just told us?

Ramon Bautista: This is Ramon again. I shouldn't need to standardize the process but is this current methodology that you propose – is this – what is the sensitivity and specificity of this methodology and how do the different measures compare to one another and picking up this particular measure?

David Seidenwurm: OK. Well, that – let me try to answer your question as I understand – as I understand it. What you're asking – what you might be asking is what is the sensitivity and specificity of each of the individual noninvasive methodologies and invasive methodologies for measuring the carotid stenosis.

Ramon Bautista: (Inaudible) compared to the gold standard which is an angiography I am assuming.

David Seidenwurm: Sure.

Ramon Bautista: And again compared to other similar measures of carotid stenosis. How does this particular measure that you are proposing stuck up?

David Seidenwurm: OK. The NASCET measurement compared to the European method as it is called is more reproducible.

Ramon Bautista: Is it more reliable though compared to a gold standard?

David Seidenwurm: (Inaudible) question is more reliable in the sense of predicting the outcome at randomized trial and more reproducible among observers if that's what you're asking. And then the different noninvasive methodologies have differing sensitivities and specificities compared with the gold standard of carotid angiography.

Ramon Bautista: (Inaudible) would a neck ultrasound pick up this measure well, you know, compared to…

David Seidenwurm: Oh, yes.

Ramon Bautista: …for example, the velocity measures it actually does right now.

David Seidenwurm: Sure. The carotid ultrasound – the Duplex ultrasound is the least accurate – if I'm aloud to use that term to combine sensitivity and specificity – it is the least accurate of the

methodologies and the least reproducible among observers. But having said that, it is certainly better to have the most reproducible gold standard against which to measure it rather than the previous hodgepodge of measurement methodologies that were employed.

Ramon Bautista: OK.

Karen Johnson: OK. So looking at impact, I think you were (inaudible) looks like it's a big problem, carotid stenosis. So I don't think anybody disagreed that the impact was high. And then performance gap, everybody agreed that it was a high performance gap because you gave some data from both the 2008 PQRS and the 2010 PQRS and basically shared what's the (inaudible) performance totally around.

So unless there's any extra points on impact for performance gap, we can go straight to the evidence. OK. So for evidence, you as a developer use the guidelines as systematic review and some additional literature. And looks like the ratings everybody agreed that there was evidence underlying this measure and there is a little bit of split between whether the quantity/quality inconsistency with high or moderate but made disagreements about that.

I guess the – I will ask either the workgroup developers or even the workgroup members or the developers maybe the same question that I ask for the epilepsy measures which is what is the specific process outcome relationship? So if you do this imaging – if the reports have this referenced what does that help? What is – how does that improve patient outcomes? What's that link?

Diedra Joseph: Hi, this is Diedra of the PCPI. I'll take that one.

Karen Johnson: OK.

Diedra Joseph: In Section 1C1, we included the link outcome such as more accurate quantification of stenosis as well as more appropriate treatment based on (inaudible) stenosis; however, the data that we included to support the measure also shows how using the standardized – the more standardized way of quantifying stenosis helps to predict the outcome of potential surgery for treatment. So hopefully that answers your question.

Karen Johnson: OK. So the standardized definition relates to better quantification which leads to more appropriate treatment which then hopefully they show better outcome. Is that correct?

Diedra Joseph: That's correct.

Karen Johnson: OK. And the evidence that you presented shows that link that having that standardized definition really does – did it I think.

Diedra Joseph: Right.

Karen Johnson: OK. Any workgroup members have any questions about the evidence or any points that you want to bring out on the evidence that is presented by the developer? There was just a

little bit in 1C14, a summary of the controversial – contradictory evidence and you talk about some potential challenges to the NASCET criteria. Is that something that you still need to be handed on a little bit or…?

Diedra Joseph: This is Diedra again. The reference that was cited there on the contradictory evidence section is just the publication that I found by Allen J. Fox. It was only one study by just this one physician but it is from 1993. We just wanted to be sure that we included all of the data.

David Seidenwurm: And I think if I may – this is David Seidenwurm again. It is important to note that that was essentially documenting the gap in care that people were incorrectly employing the methodology of the study and suggesting that they were employing the methodology. So that's kind of – what we're after here is the learning curve of – and correcting the gap in care.

Karen Johnson: OK. Thank you. So let's go into the scientific acceptability. So this measure was tested for validity and reliability. And it looks like committee members were pretty happy with the information at that time. Diedra (inaudible) said I have plenty or several questions because I have been confused about what was done for the testing.

And I still have one question, Diedra. I'm still a little bit confused when you talk about the (intraradar) reliability that was done for the EHR site, the two EHR sites that you did. And can you just tell me exactly what was compared because you mentioned that you – that you asked whether the EHR could do it with actual data elements? Was it a report? I guess I'm still confused about what was compared. So I know you looked at the entire EHR, the full HER.

Amaris Crawford: Let me see if I understand your question. I'd like to clarify. This is Amaris from the – Amaris Crawford from the AMA. So you're asking if EHR data was compared during (intraradar) reliability testing. Is that your question?

Karen Johnson: No. It is really – I think – I believe that you told us that there were two manual extractors revealing electronic data and the results were compared and that's how you plot your (intraradar) reliability numbers which were, I think, 100 percent. So I was just curious. It's hard for me to get in my head what is being looked at. I mean if you were coming from paper records I would understand and I just wanted to make sure I understand what you're doing when it is no paper records that you are looking at that when it's EHRs.

Amaris Crawford: Right. So when it's EHR information, we're looking at the – we have our what we called parallel forms reliability testing and that's different from (intraradar) reliability testing and you kind of have it right. (Intraradar) is looking at the two paper records or two obstructors taking a look at a paper chart and comparing. With parallel forms reliability testing, we're looking at the EHR, and then in this case for our testing site one, we were looking at the radiology information system, and four testing sites 2 and 3, we were looking at our EHRs and having our obstructor compare the information in those systems.

Karen Johnson: But what is compared?

Amaris Crawford: For intraradar or for parallel forms?

Karen Johnson: The…

Amaris Crawford: We're looking at the EHR versus the patient medical record.

Karen Johnson: So I want to make sure – I'm looking at page – well, you might not have the page but I'm looking at your reliability information. So you have an overall reliability, your numerator, and denominator reliability 109 records, 100 percent agreement. So again, like for site 2, just for example, did you – are you saying you got a report from the EHR and then you looked at the report and looked back at the four EHR in a match. Is that what you're saying?

Amaris Crawford: That's correct. That's correct.

Karen Johnson: OK.

Amaris Crawford: That's our parallel form.

Karen Johnson: OK. That's what you are calling parallel forms. OK. So just to clarify for the workgroup members, when you're using the full EHR as a gold standard which sounds like what – that might be what you're doing in this case. You're even going beyond what we call reliability and you're getting to what we would call validity of your data.

Amaris Crawford: EHR validity against the gold standard. I believe that's what you follow.

Karen Johnson: Yes.

Amaris Crawford: I'm sorry. I should have used that verbiage instead.

Female: OK.

Female: Instead of our PowerPoint.

Female: Yes. I think and that might be something that you can work on when we open up your forum for you. That way everybody understands the terminology. And I know we're getting very close on time. I don't see many comments about usability and feasibility and pretty much awful work rate number is – did…

Male: I did have one question about that. So how easy – I mean, you say that you're going to have direct or indirect reference to measures of fiscal and current credit. What's an indirect reference to that? I mean I'm not sure – how do you pull this information now?

Male: May I try that one again.

Male: Sure.

Male: OK. Let's take the example of – the simplest example of an indirect measurement would be to the time of flight MRA, which is a particular technique, like I say it's less used now, although I did one a couple of days ago – the only way to get pictures of this particular patient.

In that condition, with that technique, the measurement of the distal lumen is indirect in the sense that you have a pilot MRI and you have a pilot angiogram and you know, catheter angiogram and you compare them with each other but you infer from signal patterns, what the degree of stenosis was on the angiogram using the reference to the distal internal carotid diameter. So that would be indirect.

A direct measurement would be for example on the – on a catheter angiogram itself or a CT angiogram or using other MRA techniques where you actually put a cursor or a ruler or whatever you want and actually measure it physically and compare the two diameters rather than doing it inferentially using the measurement of your gold standard.

Male: So on the radiology report though, if somebody just – you know, there is irregularity in the carotid bulb, whatever that there's 60 percent stenosis but they don't specifically mention internal carotid diameter – I mean so how does that – how do you rate that?

Male: Sure. The way this would be done and the way this is done – for example, in our practice is went through all of our tables that we use for cross tabulating, for example, ultrasound got everybody – all the radiologists from train so to speak at, you know, how you actually do the measurement and we actually, you know, went through and did this and so operationally, you would specify for example the ultrasound velocities and velocity ratio and then you would calculate the stenosis and then you would state that you use – we don't – we say the (nassid) method was employed because that's what's understood by our surgeons and you know, cardiologists, neurologists, interventional radiologists and so forth.

So that's how we do it in our practice. I know there would be other ways to do it as well.

Male: But this would take somebody – I mean, you have to have an abstract or looking at the radiology report (inaudible).

Female: You could do it that way or you could do it by, you know, by attestation as other things are done, you could search for the data elements in the records – either paper record or electronic record.

Karen Johnson: Any other questions or comments about this (inaudible) measurement measure?

OK. If not, I'm going to hand it over to Suzanne to conclude the call for us.

Suzanne Theberge: OK. Thanks, Karen. (Kathy), can you open the lines and see if we have any public questions or comments?

Operator: If (inaudible) feel like to ask a question, please press star then the number one on your telephone keypad. We'll pause for just a moment to compile the Q&A roster.

And there are no questions at this time.

Female: Great. Thank you. Well, thanks very much to all the steering committee members for a great call today and thank you to the developers for your time and for answering the questions. Next steps for the committee are that – we'll see you in a couple of weeks at the in-person meeting on October 3 and 4. It's hard to believe it's already a month here but we're looking forward to seeing you then so what you'll want to do is just to start taking a look at the rest of the measures, which are also shared on SharePoint so review those in preparation for that meeting.

And for the developers, we'll follow up with you individually regarding further information in your measure form. Does anyone have any questions about next steps?

Female: This is Diedra at PCPI. Just curious, is the registration open yet for the meeting?

Female: Yes, it is and that's a great question and actually, all steering committee members and measure developers should attend – I mean to register if you're going to attend either by phone or in person. So we'll make sure that we've got everything set up for you.

Female: Thank you.

Female: Any other questions?

All right, well with that, I think we can conclude this call and if anybody does come up with any questions, please don't hesitate call or e-mail us.

Thank you.

Male: Thank you.

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

END