NATIONAL QUALITY FORUM

Moderator: Suzanne Theberge
November 16, 2015
2:00 p.m. ET

Operator: This is Conference #: 16877427.

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

Suzanne Theberge: Hello everyone, and welcome to the Pediatrics Performance Measures Workgroup Call 3.

Thank you so much for joining us today, we're really looking forward to having a conversation about the – some of the measures that we'll be looking at in this project.

Before we begin the call I'd like to, just in some general housekeeping announcement. The first is to let folks know that you have to be dialed into the phone line to speak. We can't take audio input through the computer so although we are streaming slides and audio if you wish to listen only, committee members and developers should dial in so that you can talk on the call. And we can send out the phone number via chat. So, please do dial in.

The other couple of housekeeping request are – since we are streaming audio if you are on the phone and on the computer please do turn the volume off on your computer so we don't get interference or an echo. And then also if you can keep your phone on mute when you're not speaking, that also helps with the clarity of the line and please don't put us on hold because we will get your hold music.

So, that's everything and I think we can just jump right into the call. We'll start with our introduction, let NQF staff introduce themselves and then we'll do a committee roll call. So, as I mentioned this is Suzanne Theberge, I'm the Senior Project Manager on the team. Nadine Allen, can you introduce yourself?

Nadine Allen: Hi I'm Nadine Allen. I'm the Project Manager of the Pediatrics Performance Measures Project.

Severa Chavez: Good afternoon everyone, this is Severa Chavez and I'm the Project Analyst on the team, thank you.

Robyn Nishimi: And I'm Robyn Nishimi, I'm a Senior Consultant at NQF serving as Senior Director.

Suzanne Theberge: All right, thanks everybody. OK, so now we'll do our committee roll call, if we can just jump to the next slide. And, after I call your name if you could just let us know that you're hear and give (one-tenth) of introduction about yourself.

So, next slide. Martha Bergren? Martha? OK. (Cara Burst) around?

Martha Bergren: Yes, hello, sorry they were trying to connect to me. Yes, this is Martha Bergren, I am a nurse who's been involved in measures before and was on the child health steering committee for the National Quality Forum.

Suzanne Theberge: Great, thank you. (Cara) are you here?

(Cara Burst): Yes, this is (Cara) I'm here primarily as a consumer and has done a lot of work in patient and family engagement but I'm new to the measurement strategy.

Suzanne Theberge: Great, Karen Harpster?

Karen Harpster: Hi, I'm here, I'm a Pediatric Occupational Therapist and researcher here at Cincinnati Children's Hospital.

Suzanne Theberge: Great, David Keller? Are you here?

David Keller: I hear you, coffee with an apple pie, I thought we were further – I was further down the alphabet. David Keller, Vice Chair of Clinical Affairs and Clinical Transformation at Children's Hospital Colorado, and a pediatrician.

Suzanne Theberge: Thank you. (Jill Moore Gordon).

(Jill Moore Gordon): Hi, I am a developmental behavioral pediatrician, I'm a Senior Medical Director here in Massachusetts and work for MassHealth which is the Medicaid plan.

Suzanne Theberge: Thank you. (Virginia Moyer).

(Virginia Moyer): Hi, this is (Ginny Moyer). I'm the Vice President of MFC and Quality at the American Board of Pediatrics. I'm a general pediatrician and I am possibly the only person who can't see the slides.

Suzanne Theberge: All right. We should be – we'll mostly be screen sharing the measure worksheet. So, if you can't see the slides you can pull this us if you're at a computer. But actually we'll mostly be speaking so it won't matter too much. And, Kevin Slevin.

Kevin Slevin: Hi, I'm Kevin Slevin I'm a pediatric infectious diseases physician by training but I'm also the Chief of Quality and Safety at the Joseph Sanzari Children's Hospital in Northern New Jersey.

Suzanne Theberge: All right, thank you. And we have the developers on the line as well. And – so if you have questions about the measures, they can answer those questions or provide more information. So, developers just keep an ear out and we'll ask questions as they come up.

So, next slide please. We'll just go over meeting agenda briefly. What we're going to do today is go through the five measures that you all have reviewed prior to the call, and we'll discuss your comments, we'll use the comments that you submitted on those surveys to kind of help frame the discussion. If there is an area where everyone kind of generally agreed, you know, that something was fine, we won't really spend a lot of time discussing it but we'll use those

comments to find where there are either disagreements or where people had concerns.

We'll be walking through each of the NQF criteria for the measures, for each measure but we won't be voting on this call. This is just discussion only. And, after the committee finished discussing everything we'll do NQF member and public comments a couple of minutes at the end. And – then we'll go into the next steps for you all.

So, next slide please. We – this is just a list of the measure that we're going to be looking at and we're going to start – I think actually the agenda has us starting with Measure 2817. So, perhaps we're going start with that. So we will keep up the discussion, Measure 2817, Accurate ADHD Diagnosis. Before we go into that, does anybody have any questions?

All right, Nadine, we'll turn it over to you to kick off our discussion.

Nadine Allen: Good afternoon everyone. So, the first measure that's up is Measure 2817, Accurate ADHD Diagnosis. This description of the measure, the percentage of patients ages 4 to 18 years who's diagnosis ADHD was based on clinical exam with a physician or other health care professional.

Moving on to evidence. To developer link at accurate diagnosis of ADHD to increases in appropriate treatment and decreases in appropriate treatment, which leads the patient resulting in improved quality of life, grade, and functionality. For the NQF algorithm of evidence, the eligible rates are high, moderate, or low because the developer identifies a systematic review that is graded and that assesses the quality, quantity, and consistency of the evidence.

And so I would like to at this point, there's some questions that we would like the committee to discuss. One of the question is the numerator construction is including inattention, hyperactivity, and impulsivity. Is there evidence that other symptoms should be included, does the committee wish to discuss what the developers whether all three must be present.

So I would like to open it up at this time to the workgroup members to have that discussion. I know that we also had an opportunity where we send out a

committee survey for you to provide some initial responses on the evidence. And some of the responses were very good feedback for the developers and for you fellow committee members. So, I would like to open it up at this time for discussion.

(Jill), would you like to go ahead and get started?

(Jill Moore Gordon): Sure, I have a couple of thoughts about these criteria of all three needing to be present. That gets you – part of the group of people – kids with ADHD. But does not get you all of the minutes especially if you're – sort of thinking about the DSM-IV where the, you know, they're sort of talking about (being) attentive group only, the ones that have more impulsivity and hyperactivity. So, those were just my thoughts in terms of approaching the diagnosis of ADHD.

(Virginia Moyer): This is (Ginny) and I was one of the one who was completely unclear on whether the measure required all three or some. And so, I had difficulty assessing it because I couldn't tell what was required. And then the other concern that I had was about the – how these things were documented. It appears that using a checklist, if a checklist is done by a physician or documented by a physician, is that adequate? So I actually need feedback from the developers about the specifics. As reading through the documentation, there were some variability between the word "and" and the world "or" among different descriptions so I couldn't tell for sure what was required.

Nadine Allen: The developer on the phone?

Female: Yes. Yes the developers are on the phone.

Nadine Allen: Can you respond to Dr. (Moyer's) question?

Female: It – any of the three, in the appropriate scoring would result in a accurate ADHD diagnosis using a validated tool.

(Virginia Moyer): So there was a requirement for the use of a validated tool that a physician exam documented was not sufficient, there had to be a validated tool?

Female: There had – well, you could also – it says in the measure itself, you could do a direct assessment or a validated tool. But the validated tool all go by that criteria and therefore are going to be more reliable but if a physician does go through all of the appropriate ADHD, DSM-IV criteria it – and, conducts a physical exam that also would meet the measure.

(Virginia Moyer): So, direct assessment would require documentation of each of the DSM-IV criteria?

Female: That's right, that's right. Because that's what you would also get from a validated tool. But you'd have to know that all of the appropriate assessments were done.

(Virginia Moyer): In the field testing of the tool, how frequent was it that – of the measure, how frequent was it that this was direct documentation or direct assessment rather than a...

Female: Almost never. Almost never.

David Keller: This is David Keller. I was wondering that very same thing. And, I was particularly – you're wondering because it seems to that if you're doing direct observations you would be able to meet the two location rule. Because you can't get observations from the second location without having some sort of document that's gong to gather those.

Female: You can get – you need...

(Crosstalk)

David Keller: Unless you went out and made a site visit like sat at the back of the kid's class or something like that. Is that what you were thinking?

Female: No. That the physicians would receive informations from a site that satisfied all the criteria of the DSM-IV criteria and that the clinician themselves could directly assess the child as well, and must assess all of the elements of the DSM-IV criteria.

David Keller: OK, so...

Female: In practice – so the point is, in practice, it's when the element are done they're usually done using one of the standard validated tool. There are several non-validated tools that do not include all of the elements, and that's actually a little bit – so are they assessing impairment in two settings? And are they assessing all of the elements that are all part of the DSM criteria. So, it – that's really the point.

And we made a distinction here that direct assessment if it does meet all of the criteria for the DSM definitions. Because, we felt that just because someone didn't have a validated tool in their practice, it could still work with and assess a child. So – and that's what the actual diagnostic manual says, you have to assess those things. So, we gave the opportunity for that but in truth we did not really see anybody trying to do this somewhat independently without any tools, and really it ends up being a distinction between validated tools, no tools, and non-validated tools.

(Virginia Moyer): So I was struck by the really rather low (Kappa) particularly if it was using these – we're using standardized tools.

Kevin Slevin: So, this is Kevin Slevin, can I just – to ask a questions back on the developer, if – since the DSM-IV criteria allows for certain aspects to be present, you don't – does a person making the diagnosis that we're doing it by direct observation with information coming in from outside, do they have to look at every single criteria or if the patient would meet criteria based on the information that they have without looking at all of the different criteria, would that be enough to satisfy the metric?

Female: They have to look at all of the criteria...

Kevin Slevin: OK.

Female: ... that's the requirement of the DSM-IV. And it is the requirement of the latest guideline. So, it can't be just a scan jump over sort of thing.

Kevin Slevin: OK.

(Jill Moore Gordon): So this is (Jill Moore), I'm a little confused. I'm hearing (sort of) three things, one is direct observation which I would interpret to be looking at the kid in the office. The other is meeting DMS criteria, some of which you're not going to be able to directly observe, you might get by history but you're not necessarily going to be able to see in the office. And, my confusion about direct observation is there are lot of things that potentially look like hyperactivity in one setting that may in fact not be hyperactivity and just direct observation unless I misunderstood that. I don't think (my) picked that out.

Female: So the – I think the confusion is in the word observation, we do not say observation we say direct assessment which means – which can mean talking with the parents or guardian. It can mean talking with other community members. It could be receiving information from school settings or other settings. Does that make sense? It just has to follow the DSM-IV criteria. And it has to be in two settings, so it's not just so the physician is standing there observing the patient and determining that particular context.

(Jill Moore Gordon): So they wouldn't they then be verifying the DSM – the presence of the DSM criteria in more than one setting.

Female: Correct.

(Jill Moore Gordon): OK.

Nadine Allen: OK.

Female: And I hope I answered the questions about – it can be hyperactivity, only impulsivity, only or the hyperactive impulsive type. They all are assessed through these diagnose treatment.

Nadine Allen: OK. We're moving on now.

David Keller: Can I just ask – this is David. Can I ask a procedural question to our stuff? So, one of the comments we have was that it was in – the wording through the application was inconsistent on that last point. Is this a time where we can ask

the developer to perhaps review the application and clarify that language so that we don't get hang up on it at our meeting?

Nadine Allen: Sure.

Female: Yes, yes.

Female: And by what process should we do that?

Female: Well (get back) with you after the call.

Female: OK, great.

David Keller: Thank you.

Female: Apologize for the confusions.

Nadine Allen: OK. OK, thanks. So we're moving on to performance gap. The developer provide a performance results for the measure using data abstracted from 118 charts across four outpatient and clinical office network in the Chicago area. Performance rates varied from 63.41 percent to 92.86 percent across the four sites. The developer knows that racial, ethnic disparities are found among the patient population included in the four outpatient clinical office network.

I would like to open it up for some discussion around, is there a gap in care that warrant a national performance measure or this measure be indicated as disparity sensitive? We also have some feedback from several of our committee members and some concerns about the performance gap. I would like to first ask David to provide some information about that as well as Kevin had some questions about the gaps in performance.

David Keller: I think...

Nadine Allen: David?

David Keller: Yes, this is David Keller. I didn't think I have any questions about it as much as – I thought there's a 118 charts is not a lot on which to decide if there's a national gap. I do think that there probably is substantial variation across the

country. I think this was a huge problem, and I think it will likely to, you know, if were to look for it we'll be likely to find it, it's just hard to base the national standard of, you know, four sites of the Chicago area, much of it like to think Chicago is a Microcosm of the country. It probably isn't. I think that was my only concern.

Kevin Slevin: And the concern I had – it kind of mirrors that but it's a little bit more specific in terms of the volume of patients that were included for the different racial and ethnic categories. As well as sort of disparities that have become – that are known that I'm not quite as familiar with in terms of how patients are diagnosed and the accuracy of the diagnosis among different – or misdiagnosis in different ethnic groups, in different cultural backgrounds.

Female: There's actually nothing at all that tells us about missed diagnosis, these are – the denominators is people who are diagnosed. So, if – and obviously the questions is were they – were the people who got the diagnosis, was it appropriate that they get the diagnosis, the quality gap I would think would be – whether we genuinely think that the people who did not pass the children who's documentation did not pass this quality measure, were they in fact inaccurately diagnosed, and as a result inappropriately managed. So that's the critical thing. If these are people who are being diagnosed when they shouldn't be diagnosed then we have a quality problem. If they're being appropriately diagnosed but poorly documented that's a different kind of quality problem.

David Keller: Yes, this is David Keller again, that was my thinking. And certainly anecdotally, I mean my practice is in a Medicaid population that's quite diverse. And I hear lots of anecdotes about children of color who are inappropriately diagnosed and medicated for a variety of reasons and feel that (this stuff) has not served them well.

So, finding this – even if in the small samples that you have here, I know there's degree of variation that you found, doesn't surprise me and makes me feel like there really is a problem here that we need to address. I was a little also concern that the race ethnicity that was unknown was, as high as it was

but of course that's a common problem anytime we try to look at race ethnicity in any kind of sample.

Female: Would it be helpful for me to say more about that? So, this measure is based on this notion that came through the AAP ADHD guideline. That the appropriate diagnosis of ADHD must be done according to the DSM criteria which means either using a validated tool or, as we just discussed direct assessment. What we found in terms of a disparity particularly sharply was that validated tools were not used and appropriate diagnostic processes were not applied in racial – in a greater proportion, significantly greater proportion of minority populations.

Nadine Allen: OK. Moving on to reliability. The measure specified that the group practice and facility level of analysis, it uses data from either electronic health records or paper medical records. This measure captures children ages 4 to 18 years with a diagnosis of ADHD. There are no exclusions specified for this measure. The developer provided ICD-9 and ICD-10 (close) that identify patients diagnosed with ADHD. An assessment of core symptoms can be dug through use of validated instrument or through direct assessment.

Developer also encourages stratification of measure result according to gender, age, group, race, language, and insurance type whoever this stratification is optional and is not meant to be served as a method of risk adjustment. So some questions for the committee that staff had was, are there instruments cited as examples in the submission appropriate for assessing the core symptom as ADHD? Are all the data elements clearly define? Are there appropriate codes included? Is the logic or calculation algorithm clear? Is it likely this measure can be consistently implemented?

We also had some feedback from a couple of our committee members. I would like to first ask (Virginia) to talk a little bit about our comments around the reliability of specification.

(Virginia Moyer): Well the primary concern that I had and I do have a question for the developer because the – some of the information that we received, it said that the

measure was specified at the facility level. But I didn't see that it was tested at the facility level, is that correct?

Female: So we worked with networks in the Chicago area. So, the primary care network of Stroger Cook County, the primary care network of Lurie Children's Hospital and the primary care networks of both advocate (Lutheran) general in the suburban environment and Christ Hope in a suburban environment. So, if we try to, you know, we were looking across their primary care networks. And so, we aggregated the data by network I guess is a better term.

(Virginia Moyer): Yes, the question that I have actually is – what's mentioned is inpatient psychiatric settings but I don't see anything about inpatient psychiatric settings. Was it intended for use in that setting?

Female: Our expert – technical panel said that it could be used in an inpatient setting but it was assessed generally at the outpatient. It could be applied in an inpatient setting. But, I think it's a plausibility in an inpatient setting is less appropriate, or (less helpful), I guess that would be it.

(Virginia Moyer): And so, what I saw that was concerning was that (Kappa) and I'm wondering if you have information on where the disagreement was that resulted in – I mean these are pretty low (Kappas) for something that were – if you're saying that primarily people were using standard tools. Are they just not using the tool or?

Female: We were surprised by that as well. I don't have a lot of guidance about that.

(Virginia Moyer): OK.

David Keller: And this David Keller again. So my question was a little bit more basic because the diagnosis of ADHD, anywhere in the start, if you have a child who you've been following for 10 years is a pretty broad definition. There's actually children who can be diagnosed with ADHD at age six who sometimes – you discover two years later you were wrong. And yet, the diagnosis is in the chart. So, is that patient part of you denominator in that case?

Female: No. Yes.

David Keller: So how are we...

Female: It will be in the last year.

(Virginia Moyer): So – but, they have to have been diagnosed in the last year or it has to have appeared on their problem list in the last year?

Female: In diagnosed in the last year.

(Virginia Moyer): So these are all newly diagnosed patients effectively?

Female: Correct.

Female: Correct, we're assessing, are they accurately – getting accurately diagnosed? So we're looking at the diagnostic process.

(Virginia Moyer): You're looking at the diagnostic visit, not at any other visit?

Female: Right.

David Keller: OK, but still there maybe patients who have multiple diagnostic visits for multiple reasons overtime. And, actually there's sometimes disagreement among different petitioners. So the other group I was wondering about are patients who are diagnosed at another – by another provider (inherent) for some reason. That they are coming to you because filling your (chart) in your charge it's going to be a new diagnosis, the first they come to you – but they were actually diagnosed two years ago by someone else.

And for that patient I might not have – I might not go through the same diagnostic workup that I would go through for a different patient. And yet I would accept the diagnosis of – I might accept the diagnosis of a colleague and feel that I don't need to go through that process again to clarify that diagnosis. Does that – I'm not sure I'm being clear but does that make sense? I mean (if you have to do) transfers into your practice that has diagnosed previously, it would seem to me you want to exclude that child but you don't really have a good mechanism for that in this measure.

Female: It's a little more complicated than that, in that the current recommendation, I mean we have this in a – our subsequent measure which is the chronic care measure. The guideline recommends that patients be seen within – three times within the first year of a diagnosis to be provided appropriately treatment in order to assess the medication treatment or behavior therapy treatment. And also to treat these patients that – chronic care patients and that they should be reassessed.

So, we would – we understand, we know that the state of care right now doesn't actually incorporate those recommendations. But, that should help as those guidelines become more standard of care.

David Keller: But getting back to the – my question really is that – is a reassessment of things like – so in your estimation or reassessment by a new provided would be the same as a diagnostic visit for a patient never been diagnosed before? Holding the same standard?

Female: Actually that's a very good question because, really what we're most interested in is finding out if – for this measure, if a patient is diagnosed appropriately. And, we believe that the science behind the DSM – or if we are to believe the science behind the DSM-IV criteria, then they will be inappropriate diagnosis made.

We're – I guess I agree with you, it's really where most interested in the diagnostic visit. With appropriate treatment we should see a reduction in the symptoms and impairment. And that's been demonstrated as well. So, reassessment is actually to assess whether the treatment is effective.

David Keller: Right.

Female: So.

David Keller: Right. But – I guess I'm wondering whether you should be thinking about some kind of exclusion for people who have been previously diagnosed.

Female: I think you're probably right. I think the measure would be improved by that.

David Keller: No further questions your honor. But thank you, no – good, thank you every much for that.

(Jill Moore Gordon): So this is (Jill Moore) and I just want to throw a little monkey (rent) in this. I know you used DSM-IV and DSM-IV-TR, but the DSM-V is out and has been for a couple of years. And, sort of – things will move on, and I guess the questions would be, in terms of the measure, it's the DSM-IV, DSM-IV-TR, would that – how would that translate into DSM-V as people move on in time?

Female: I am not sure, did not look at that..

But I can respond after this call with regard to those. There was discussion about it but I just don't have it in front of me what people have said. So, I will make a note of that.

Nadine Allen: Kevin, you wanted to say something, I see your hands up?

Kevin Slevin: Yes, actually I had a question sort of in terms of the – just in terms of the algorithm for data collection. And you will have to forgive me this is not my primary practice area. But, I imagine like many diagnosis especially of chronic condition, information comes in that gets you to the diagnosis overtime rather that it all being addressed in a single office visit. So do the numerator specifications allow for a period of look back within the chart from the time of diagnosis to how far back you're allowed to go to gather those other pieces of information, particularly the ones about impairment in areas outside of a person's specific practice arena?

So if I do direct assessment in my office but there's issues related to the child's functional impairment at school or functional impairment in social group. How far back are we allowed to go to sort of make sure that at the time of diagnosis these assessments were done?

Female: It's generally done in one office visit where there's been information collected. So, there's concern potentially brought up by a parent or a school that gets discussed in the office, then the assessment begins from there and once all of the information is in, that's when the diagnostics and termination...

(Off-Mic)

(Virginia Moyer): But when the person is assessing the chart, the abstract or who's going to be looking at this, how far back might they, you know, I can see exactly what Kevin is saying, you know, the counter rating (skill) gets mailed back in and gets chunk into the chart, and these are not paper charts, they're electronic charts so stuff is hidden everywhere.

Female: Right.

(Virginia Moyer): And how do you – does it need to be redocumented on a particular day or, how is that handled, it just strikes me as a difficult, a potentially difficult data collection that might even explain some of the (low cap) is.

Female: There isn't a look back period, there's a date of a visit that is also to include a physical exam and for an accurate diagnosis. And, generally all of the information is compatible. It is a...

(Virginia Moyer): I can easily imagine the situation that I think may not even beyond common where a child comes in with a parent, you know, concern, the teacher's concerned and you do a physical exam and you talk to the family and you document all the concerns and you document the physical exam. And then you send the parent home with a counter rating (skill), the teachers and so forth and so on.

And a month later everybody comes back and you got the skills and – but you don't repeat a physical exam you just did that.

David Keller: Right and that happens...

(Virginia Moyer): (Even) if the visit doesn't have they physical exam, but the physical exam is not relevant on that data, it was done a month ago.

Female: As part of the screening, the initial screening.

(Virginia Moyer): Right. And you make these diagnoses over a period of several visits oftentimes.

David Keller: Right, particularly if you're doing it – this is David Keller again. And thank you for – that was exactly what I used to do in my practice, I never did one of these in a single visit, it was always two visits. I did it the opposite way, I focused the first visit on getting history and spent most of the time doing – getting history in my physical exam was very cursory, it was – yet the child looks like he's fine.

And then sent him home with (Vanderbilt) and then they come back at two-three weeks later, at which point they would do the physical exam and then also talk about treatment options and other stuff. There was too much information for the parents in one visit. So I'm thinking to myself that would show up.

Female: So, that I think the way that things are generally done because the physical exam is hardly about medication issues that could emerge. So, it's frequently – I mean that's what our experience is. The physical exam is done at the time of diagnosis.

(Virginia Moyer): Yes, but it doesn't – what I'm saying that you would have – it doesn't matter what's usually done, what matters is what needs to be done to assure that the child gets an accurate diagnosis. And if we're saying everybody has to do the physical exam on this date or you have to, you know, the question is just whether the measure itself is measuring what you wanted to measure. It's a...

David Keller: Right.

(Virginia Moyer): ... (it's the) provider, document all of the appropriate things. And, the concern that I have is that the possibility that the construction of the measure and such that it's actually missing appropriately managed children.

David Keller: Well, and I'll just throw out the other piece is that both of those visits would carry the diagnosis of ADHD.

(Virginia Moyer): Yes right, exactly.

David Keller: And my chart – so if you went back and reviewed my chart you might fold the first visit and think that I've done an inadequate job because I didn't do a physical exam of you pulled a second visit and discovered I haven't taken a comprehensive history because I don't take a comprehensive history in the second go around.

(Virginia Moyer): Right.

David Keller: And, it's really those two visits together that constitute my evaluation.

(Virginia Moyer): Right.

Female: So you would recommend a one month look back period just for the insurance?

David Keller: I think we need to know what the standard of care is around the country, I'm not sure if it's – I mean that was (my) (one guy).

Female: That is what – what you describe is what's standard of care, if there's an initial screening and then information – parents are sent out with asking for information and then within weeks of a month or so people came back to the office with the information, do the physical exam and discuss treatment option.

David Keller: Yes.

(Crosstalk)

Female: The question is whether it's poor care to only have one of those visits be have all the documentation or whether appropriately you might have several visits that – where the documentation is accumulative.

Kevin Slevin: Right, and that actually model is pretty true for most chronic diseases or chronic conditions where it may take time to gather all the information, to make the diagnosis. But, just because of the vagaries of coding, you may start coding for that diagnosis with the first visit, you know...

(Crosstalk)

... even if that's not what they have.

Female: Right.

Female: Right, there's the initial screening and then there's the diagnosis. And that is understood I believe in this measure. The initial screening and the diagnosis.

Female: So I've just been digging through what we have available to us and maybe one of the (staff) people can clarify this for me. But not having the – what I have under data – what's in under data collection instrument just looks like the data that were selected, you know, kick marks or data that were selective but I don't see – what I don't see is the – anything that would clarify whether this was a particular visit or whether this over a period of a given period of time.

I don't see anything that would make it so that it would be crystal clear what constituted meeting these criteria for someone who is – actually is something the visit.

So, it would just be helpful if we knew that.

David Keller: Right.

(Crosstalk)

Female: Go ahead.

David Keller: Please keep in mind we all think this is an important measure that's why we're working on it. So – because this is important.

Female: So, if I'm understanding you correctly, it's not clear in the measure that you can look back from the day that the diagnosis and there should be A, explicit language saying, look back...

Female: Or look forward.

(Crosstalk)

Kevin Slevin: ... if the diagnosis is made now, do I identify a new diagnosis now but it's been – it's not documented until six months later, that might be problematic. But, a month or a couple of weeks to make sure all of the assessment tools are back maybe reasonable.

Female: OK.

David Keller: Right. It's – yes, I mean it's, you know, there are the clinics that will do the three hour massive assessment where you do everything all at once. But I think it's spread out – it's more spread out over time most of the time.

Kevin Slevin: Especially in a primary care setting.

David Keller: Yes.

Female: Yes. So, the model of doing it all at once is the developmental pediatric model.

Female: And there aren't enough of those to be able to do this.

Female: There are absolutely not enough of those to do that.

David Keller: That's correct. That's why we all end up doing it.

Female: We're glad you do actually.

(Rober): Karen, this is (Rober), did the committee already jump right into reliability testing. I'm not going to provide an overview about that but there was one thing that Karen brought up on the survey response and she talked about the denominator testing that there were no information for that. And several other committee members have the same comment. Would you like to speak about that at this time Karen?

Karen Harpster: I guess I was just wondering if there's, you know, specific information that...

(Off-Mic)

... on the denominator.

Female: I'm clear about what the questions is.

Karen Harpster: There wasn't much testing, are information on the testing as a denominator that we – from the overview that we were given?

David Keller: Right. How did you pull it?

Female: We identified 96110 code through an (EZW), and then evaluated those charts, looking for the ADHD diagnosis.

Karen Harpster: And did you consider the 96127 which is – it's described as the brief behavioral assessment but when you look at the more detailed description it includes doing (conners) and that sort of thing for ADHD.

Female: Which is not, we were advised by the American Economy of Pediatrics person who work with coding that it is very unreliably used and so we did not use that (approach).

And it's a – 96110 code also was unreliably used but it is reliably used within state, it's...

(Off-Mic)

So we were able to in this environment identify the diagnosed children with that approach.

(Crosstalk)

... with the ADHD diagnosis. It just – it narrow the poll for us, that's all. It's not required but it narrowed the poll for us.

David Keller: Yes it's interesting because I agree with you, Colorado we don't use the 27 code.

(Jill Moore Gordon): And here in Massachusetts we use the – it gets mixed up with the 96110 as well. So...

David Keller: Interesting.

(Jill Moore Gordon): ... kind of a mess.

David Keller: Yes, I mean we were just – we were trying to narrow the poll for selection, it's not a requirement if you look for the ADHD diagnosis.

Nadine Allen: Moving on to validity and validity testing. For validity specification the developer noted that the AAP ADHD guideline knows that, to make a diagnosis at ADHD the primary care clinician should determine that DSM-IV criteria have been meet include documentation that's impairment in more than one major set in, the major specification corresponds and require two or more setting. The criteria and the guideline specify academic of behavioral problem and symptoms of inattention, hyperactivity and impulsivity. The measure specification corresponds with this.

For the validity testing the developer did not conduct empirical validity testing for this measure. The developer states that they used face validity, they conducted this with 25-member expert panel that help develop the measure and they agree that the measure can be use to distinguish good and poor quality care. The face validity was assessed via a 21-day pubic commenting period. Question for the committee, are the specification consistent with the evidence since the DSM-IV criteria for diagnosing ADHD are sited, do they include assessment in attention our productivity and impulsivity. And do you agree that this core from this measure specify is an indicator of quality?

I would like to open it up for discussion with the workgroup members at this time. Karen you mentioned the – you had some concerns about the face and about the missing details there and also there were no formal empirical testing that was completed.

Karen Harpster: Well just had questions about how, you know, it was reviewed – the face validity was reviewed by an expert panel rather than empirical testing. And I didn't know if that was a standard of care or was there – is there a different way that that could be done in order...

(Off-Mic)

... the information. And that it was done in a small, you know, small area. And is that...

(Off-Mic)

... to the broader public?

David Keller: Yes, this is David, I share you concern particularly around, you know, the Chicago, you know, do we need to – is it likely that this is going to be as valid outside of Chicago? And – I think if we're...

(Off-Mic)

Female: The application of the DSM criteria should be standard across the country, the guidelines from the APP guideline from which this measure is generated to be extended across the country. And what we would see is variation in quality as oppose to variation in standard.

Female: I think this – I thought that the standard was relatively clear because there is not – there's actually not – the gold standard is this. This – it's a consensus standard. There are – and there's not – I mean the – there's not an evidence base that said that this is – this particular disorder, we have agreed that that's what it is.

(Jill Moore Gordon): Yes, and unfortunately this is one of those behavioral things that we don't yet have some more objective tests for i.e. a blood test of a brain MRI or whatever.

Female: I did not – I do not have an overwhelming concern about that because of the nature of the – for the disorder. I guess the question that the committee will have to ask ourselves is whether we are comfortable with the quality measure that's based on the consensus guideline. I think you're – the expert panel and everybody else, it sounds like everybody pretty agree that that's what this is.

David Keller: Right. And I did like that you have public comment. I mean you did a lot of things to bring in outside opinions outside the Chicago area. So, you know, I

wasn't too worried about the validity, I was more concerned about the reproducibility and...

(Off-Mic)

Female: Yes.

Female: Thank you. OK, Nadine you want to move on?

Nadine Allen: Yes. So moving on for (the rest) of validity, there are no exclusions for this measure, the measure is not risk adjusted, the (rates) across – for meaningful differences the (rate) across sites varies substantially and the competence into (Roche) was site one and site four did not overlap for missing data, the developer provided that the – the developer states that the denominator criteria were missing in approximately 5 percent of cases. And numerator – a criteria were missing in 34 percent of cases. The developer knows that if data are missing for the numerator element, the measure is calculated as not being met.

(Virginia Moyer): How do you miss denominator data if that's how you get your sample?

Female: I think that that was related to the 96110 method for initial selection.

(Virginia Moyer): So what happened with those?

Female: Those were excluded. Because they do not have an ADHD diagnosis documented.

(Virginia Moyer): Do you – so the denominator was based first on locating the charts with the code and then confirming that there was a diagnosis someplace?

Female: Correct.

But as I said, we did that as a method to just make things easier for us, it's not a requirement to the measurement. Because – and it's not a requirement to the measure and you'll see that in there because of the (bigger risk) of the 96110 code (cost). It's just a helpful way to identify appropriate (chart).

Nadine Allen: OK, moving on.

(Virginia Moyer): Are we still talking about reliability testing?

Nadine Allen: We've moved on, we're talking about validity.

(Virginia Moyer): Validity testing?

Nadine Allen: Yes.

(Virginia Moyer): Yes, I think the other point we need to make sure we note in reliability testing is (with each) product before, the issue about the low (Kappas) and that's using well trained abstractors.

Nadine Allen: Thanks for that, we've noted that as well. So, we can move on now to feasibility if no one on the line has any objection.

OK, so for feasibility (data source) for this measure is paper or medical record or electronic medical record. The developer reported that there are improvements that could facilitate selection and reporting of this measure, such as the documentation of the use of specific validated tools and the setting in which the evaluation took place. There's some questions here for the committee, are the required data elements routinely generated and used during care delivery, all likely (it did) that the required data element are available in structured fields and EHR, is the data collection strategy ready to be put into operational use.

There were several questions and concerns from the committee about the feasibility of this measure. I would like to open the call now for discussion between the committee members.

(Jill Moore Gordon): So this is (Jill), as to whether or not the characteristics of ADHD are part of a general exam I think they're probably not. And, from that vantage point you would either have to have something especial coded into your record or – I mean you're going to have to look at a chart. It's not going to be something where you're going to get the check boxes. And if you're making sure it's DSM – consistent with DSM, there are multiple categories under each characteristic, so under (inattention), under impulsivity, under hyperactivity.

There are multiple characteristics that would have to be identified. So I think you're probably going to end up doing chart reviews which – it takes a lot of time.

(Virginia Moyer): Yes, that's clearly what was done in the field testing...

(Off-Mic)

Female: So, one thing to know, we found that ambulatory care pediatric practices have less evolved electronic medical records than those in hospitals. We found that two of the sites that we sampled did not have electronic records at all. So, the defusion curve for electronic medical records is a bit slower in pediatric practice. And, I think there are number of groups and initiatives that are working on that now to make the (Charles Memorial) responsive to the things that are needed to be collected in EHR and things like that.

We just made those comments because, at the end of the day, you know, we're all marching towards meaningful use. And, in order for these measures to be reported at eMeasures which is a much easier method for monitoring, there would need be opportunities to document in discreet (for able fields) the elements of this measure. And there would need to be use of electronic records in pediatric (practice).

So until there's a complete evolution, we actually did test for feasibility of eMeasure – we initially had targeted developing this measure as an eMeasure but did not pass feasibility for that. But it did pass feasibility for construction in the medical, as a manual (chartered) measure.

Nadine Allen: OK. Moving onto the usability and use if no one had any – if there's workgroup members does not have any objection.

So for usability and use, the measure is currently use by the American Board of Pediatrists, for both pediatrician Performance Improvement, and as a requirement for professional recertification.

Developer did not describe how it plans to use the measure for payment and public reporting. Developer stated that there is no unintended consequences.

Would the committee like to talk about some of the question that staff had at this time? Can the performance was out to be use for the goals of high quality? (Is patient) healthcare is a measure appropriate for accountability purposes or/and do the benefits of the measure out way any potential unintended consequences?

David Keller: I was just leaping through the – spreadsheet to remember what I said about this. But my gut reaction, I'm looking it is – yes. I mean, I confess. I look at this and said, "I do think that there is substantial overtreatment for kids who are inappropriately diagnosed with ADHD. And I think this Medicaid, this measure would go a long way towards helping us moves towards, improving that. With that having been said, I think we've raise a number of questions about whether this measures actually measuring what we wanted the measure in its current form. And that's going to be an interesting point, discussion going forward.

And I have trouble, imagining unintended consequences too, but I guess that's the whole point of unintended, isn't it, is what we can't imagine them sometimes.

So I was happy to hear that on the first 118 charts, they didn't see anything, or 136 charts. But – that something we're going to have to watch for, I just can't think of any myself.

(Jill Moore Gordon): So this is (Jill). I wanted to sort of comment on one of the comments that somebody else made under this category. And, you know, in the way I was trained as a developmental pediatrician is that, it's not just sort of history direct assessment or checklist that is a matter of putting the two together to paint the whole picture.

And so, in terms of concerns and concerns for harm, you know, I think, having done this for many, many, many years, there are teachers who don't like kids and we'll do a checklist that doesn't fit the kid. There are parents who have agendas. There are kids who have other things that mimic ADHD. And I have a little concern about sort of looking at – sort of direct assessment

or checklist when in fact, what I found is sometimes you have to put them all together to make the whole picture.

That's exactly the imperilment in two settings?

Nadine Allen: No. That's using both a standardize checklist and direct assessment to put it together in two setting. So we always had the parents and the teachers do checklist. And you did a history.

And we also – I often got a note from the teacher too, which was sort of a history. So more than just – it's more than just doing it in two settings, it's even like comparing checklist between the parents and the teachers. But sometimes, you see a problem in school and not a problem at home.

And then, you have to think, you know, is it – and then, when you did the history, the kid had all the symptoms but the parents manages it at home.

David Keller: Right.

Nadine Allen: Or, they had it only in school and not at home at all. And then the question is, is this a learning disability, is this social anxiety, is this something else? And I worry that isn't something else is going to get missed by sort of either/or.

David Keller: Yes, I echo that. I absolutely agree and think that the – I'm wondering, what I agree with you that there's not a good way to measure kind of how do you integrate all this information begin a holistic view of the child, which is really what we're trying to do.

Nadine Allen: You know, you can really – so you can only really measure the pieces, you know, did you do a history that had those questions asked? Did you do the checklist in two settings? And did you take that information and measure it against the DSM?

I mean, I used to have a DMS checklist, with the characteristics and I go through those check, check, check, check, and then count them up. And, you know, and then using – and then you still have to think about it.

(Crosstalk)

Female: Yes. We're ready to move to the next measure.

(Cara Burst): This is (Cara), just go quickly from, I guess, the parent perspective especially with having – had a child who went through a long diagnostic process to come up with some with some – three integration issues and having had teachers who didn't want to use on trainer (inaudible). And we're very quite to push us back to the doctor to get diagnose with ADHD, do we put on medication. I just wanted to – I still appreciate this conversation and want to emphasize just the importance of putting value on the holistic approach and thorough thoughtful approach to diagnosis. And, you know, making diagnosis easier, it's not always better.

Female: The intention of the measure, and – which is follows the guideline, and the intention of the guidelines is to, move the field towards a more rigorous assessment of the DSM four to five criteria. From a kind of over, you know, from a non-standarized approach, that's the idea of the measure. And in fact, in some of our materials, I don't know if this is something that we submitted to you at NQF.

We describe that we think that the people maybe being diagnose with ADHD who do not have ADHD because they are – they may have some features but there's not a systematic assessment and we're very concerned about the excess use of medication in that context.

So from our perspective, that this measure isn't – is trying to prevent that unintended diagnostic consequence.

Nadine Allen: Thank you all for your comments. We're going to go ahead and move on to the next measure 2818, ADHD Chronic Care Followup.

The description of the measure, the percentage of patient, ages 4 to 18 years, with a primary or secondary diagnosis of ADHD, and the year prior to the measurement year, who have at least one followup visit in the measurement with ADHD as a primary diagnosis.

For evidence, the developer link followup visits for those with ADHD with increase treatment and ultimately improvement in function, quality of life decrease symptom, developer reported at body of (inaudible) underline in the clinical practice guidelines include three literature review, and one systematic review.

The questions for the committee, is the evidence concern in the chronic care model and medical home model, directly applicable to the measure focus? The measure specifies at last one followup visit in the measurement year, is there evidence the support, the frequency stated in specification as improving outcome.

So with that said, I would like to open it up for discussion. I know Kevin, you mention why only one visit is that enough?

Kevin Slevin: Right. And again, this is – I've always going to (preface) any of these things because it's outside of my practice scope. But, you know, over the course of a single year, a number of changes occur in the life of a child. They're in school. They're not in school. They're in camp. They're in outside classes. There's all the other things that are going on and for many chronic conditions, a single visit over the course of the year for followup to make sure that the current, therapy is appropriate for the setting that the child is currently is just not enough.

And so, the questions I just need to throw out is, is a single visit for this type of a condition enough to (warrant) a marker of quality ongoing care?

David Keller: So, this is David. I mean, I'll just point out, give you someone who (inaudible) this in, in their practice goes. We do this all the time. And I think the, you know, part of the issue is that there is – it's hard to measure followup, because a lot of the followup for this kind of chronic illness, can be done without a visit.

And I had a number – I had in my own practice, I have a number of patients who I would do telephone calls with several times a year. And we'd see them once a year. And who were stable enough requiring endorsements in their

medications. And if they needed adjustments in their medications, I bring them in for a visit.

There was a practical matter. We just had run out of slots. And so, that's how we were doing it.

So, I think there's a fair bit of variation around the country about how people manage the followup. And I know some practitioners who are very aggressive with telephone followup for this condition. And I actually think that's not a bad thing.

Kevin Slevin: So that then raises the second question, because I believe that the measure requires it to be the primary reason for the visit. But is it something if you're doing that kind of aggressive telephone conversation – telephone followup cannot be build in to the annual well-child visit that many parents...

Male: It's going to be...

Kevin Slevin: ... bring their kids into but then it wouldn't be the primary reason for the visit.

(Crosstalk)

Female: Yes. I had exactly the same concerned to have a...

David Keller: Yes.

Female: We have to – the inconvenience to parents are being required to comeback for a single purpose just so you confess a quality measure is not good.

(Crosstalk)

(Jill Moore Gordon): So, this is (Jill). And I'm just going to approach this from the developmental pediatric standpoint. When I had (to take my) medication, I actually saw them physically. And I had that very much (based) for either. But at least the couple of times a year in addition to, you know, there well visit to make sure that they were growing and didn't have high blood pressure and that sort of thing.

I do think though there's variability in terms of followup and part of that is because some of the kids are on medication. They don't tolerate medication. They have other counter indications to it.

And for them to be able to meet, I mean, I think that they do need to be seen at least once in the middle between there well-child visits unless – I guess, unless you're doing it by phone. But I don't know. There's a lot that you see actually see in the kids. And I don't know. I didn't think that there - that that's helpful.

(Crosstalk)

Female: Sorry.

Female: That is what our expert technical panel said, and related to the 2011 ADHD AAP guideline, that children should be seen in person. And have a physical exam related to their medication both to make sure that there aren't problematic side effects or other conditions that could affect there medication used or to be effected by their medication used.

And also, to assist adherence because there's literature that suggest that within two prescription cycles, most children are no longer on their initial, no longer adherence with their medication. So, on both scores, it would highly recommend that – and this is not the first year.

So, there are usually three visits in the first year to (tight treat) medication and assess whether a child is, you know, getting on a mediation that is working for them and then on appropriate dose. This is the chronic care followup. And there's also evidence about dose reduction if in appropriate course of treatment is followed.

There's a reduction of symptoms and impairment for ADHD. There's actual systematic evidence out there regarding that. And that – when children go off their treatment, whether it's behavioral therapy or medication, they often return to baseline, so it does not – it's not a curative, the treatments we have are not curative. They just addressed symptoms and impairment.

Kevin Slevin: May I ask some other question that related to this? It's a lot of the things I heard about like in-person visits related to blood pressure management, you know, weight gain or other side effects that you might – that for somebody who's on medication.

If the child is being managed without medication, is the in-person visit necessary if this therapy is working or cannot therapy be (tweak) over the phone with discussions with the parents without bringing them in. And then is this measure really appropriate in that situation? Again...

Female: Though, it's still considered – the AFP did not make the four and five-year-old a not a chromic condition. They would – the recommendations suggest behavior therapy for four and five-year-old who are diagnosed as the first line of treatment.

But they did not make any special designation that they do not have of chronic illness. And so, we believe that based on what the recommendations are that a child should be seen to assess symptoms and impairment if they are not receiving medication treatment.

(Crosstalk)

Kevin Slevin: On the guideline, I think there are some things that I'm not (running yet). So, you said there was – this is in the second year, but I couldn't tell whether or this – how do you calculate the years again without having the measure...

(Crosstalk)

Kevin Slevin: ... how do you calculate when the year begins during which year evaluating this?

Female: You have a look back period for the measure.

Kevin Slevin: From what?

Female: From when you're assessing. So – for the claim...

Kevin Slevin: So, the index time frame it somebody makes – somebody assesses the claim, OK. If there's a claim, then they were seen. So, how are we looking at – are we looking a newly diagnosed children in an given time frame?

Female: No, no.

Kevin Slevin: How are we finding these kids?

Female: OK. So, this is an entirely different measure. We are not looking newly diagnosed patient and we have a look back period over a year for that very reason, because one needs to be seen more often in the first year of the diagnoses.

So, pull up the, again, the information on – look back period and how it actually worked. Because you want to – this is intended to look at the chronic hear of the patient. This is – they have been diagnosed by some method, hopefully, a good quality, and determined to have ADHD. There's a period of time where medication or other treatment need to be – medication is (inaudible) other treatment assess potentially.

And then what this measure is about is that, the AAP guidelines have said, "This condition should be treated as the chronic condition and should therefore be seen at least once a year. The child should (inaudible) at least once a year.

Kevin Slevin: Right. So, I'm just trying to figure out how you and have – how I would, if I had a chart in my hand, how I would tell if the measure had been met?

Female: This is a claim space measure.

Kevin Slevin: OK.

Female: So...

Kevin Slevin: So, it's says – the (inaudible) have a percentage of patients. They portrait 10 years of the primary, secondary, diagnosed of (inaudible) desk that hyper activity disorder in the year prior to the measurement year. What is the measurement year? Is it a calendar year?

Female: Yes.

Kevin Slevin: OK.

Female: So, you'd look at this year and they would have – to be included, they would've had to have at least a ADHD diagnosis in a previous year.

Kevin Slevin: OK. So, let's say we're doing this in January. So, we look at 2015 and we looked then – and we find a whole bunch of child – so then, now, we need to look back to 2014 and our baseline population will be children who were diagnosed in 2014.

Female: That's right.

Kevin Slevin: And then the measurement year will be 2015. And we're taking as our baseline, the kids that were diagnosed not who already had it, but who were diagnosed.

Female: That's right. So...

Kevin Slevin: Is it right?

(Crosstalk)

Female: It's an 18-month look back period, actually.

Female: I have to get the actual – we did a lot of work on this look back period, at least to get the information.

Kevin Slevin: OK. So that would just – and I don't think we had (settled) that right now. But that would be helpful to have. And then, the other concern that I have is that followup visit has to have ADHD at the primary diagnosis. So, if you have the unfortunate child who had booth asthma and ADHD, you have to make separate appointments for followup for each one of those.

Female: There – maybe you can state to this better than I can. But the focus of the visit has to be about ADHD. You have to be assessing the child for...

Kevin Slevin: So, if that child were coming in for followup of both his asthma and his ADHD at one visit. That if the – then you could not – that would not count, because that would not be the prime if asthma. It would just depend on who happened to code the chart, which one went first?

David Keller: And I put to you that that's an incredibly common situation in private practice.

Kevin Slevin: Absolutely.

David Keller: We're not going to get the – we're not going to get them to come in for separate visits for those things. And to the...

Kevin Slevin: Well, that's unfair to the family and the child.

David Keller: Right. And so – and to their earlier point, we will also wrap both of those followups up into a widest visit we can.

So, if and again, depending on the coding rules, I'm sometimes told that I shouldn't bill clinical diagnosis going to billing a well code because it screws up decoding. And sometimes, I'm told I should put all of the diagnosis then. I've been told so many different things by so many different coders. I can't keep it straight anymore.

Kevin Slevin: Great. I think this is another one where the intense of measure is clearly, extremely positive. The concern I have is in the specification. Kids with chronic illness ought to be appropriately followed up.

Male: Oh, I completely agree.

Kevin Slevin: Right. But what we don't want to do is to develop a situation in which a provider says to a parent, "Well, I'm sorry I can't talk to that today. You'll have to come back tomorrow or in (advance).

Female: Right, right.

David Keller: Right.

Female: So, form your perspective, since we really do want to assure the chronic care for children diagnosed of ADHD.

In you're experience, would it be appropriate to look across all of the denominator and CPT code? So that, you know, the asthma assessment as well as the ADHD assessment if there was...

Male: Yes.

Female: ... and ADHD assessment listed as...

David Keller: So, this is where you get into the problem of using a claims-based initiative.

On one of the other questions, somebody asked is why is this only claims-based, why is this also a (chart review) with you. And I certainly understand claims-based assessments are in some ways simpler, because you're searching databases instead of having, you know, chart reviewers going to chart.

But it's really one of those once – if you want to get down to that level of granularity. You'd honestly have to look at the note. I mean, if somebody is coding both ADHD and asthma in the same visit, you have to believe both were addressed. And, you know, if you want to know how they were addressed. I don't see how you can pull that other claims data. I think that's when you have to get into a chart review sort of measure.

Female: So the reason we left with the claims-based measure is because of our experience with the PQMP program of talking with a lot of the Medicaid offices in the states who are generally unwilling to consider anything that has chart review related to it.

David Keller: Yes.

Female: And because we wanted – we knew it was – we initially tried to build that accurate diagnosis with the claim. And we run into all of the coding concerns that we discussed earlier. So we moved the way from it completely.

But in terms of the chronic care, there are – there is an ability to identify these children and whether they are seen. I – what I'm hearing from you is that, our

boundaries might be a little bit too tight and we may need to assess any notation of ADHD. Not just the primary codes.

Robyn Nishimi: Well, I, you know, obviously what you would like is to have a good followup visit, but if you're – what you don't want to do is develop a perverse incentive...

Severa Chavez: Right.

Robyn Nishimi: ... to have people break visits up and have people comeback more often. That's not good for family.

Severa Chavez: I think we have discussed pretty thoroughly some concerns that the workgroup has with the specification. Is there anything the workgroup wants to comment on – about the actual reliability or validity testing at this point? And if not, we'll move on to feasibility and usability and use.

Robyn Nishimi: I had some concerns about what the exclusion system reasons whether this would be something that could be implemented consistently?

Severa Chavez: OK, great. Anyone else?

David Keller: And the other concerned, I'm trying to find it but I remember having some concerns about the codes that they were using because some of the codes that were acceptable for follow up were inpatient codes or mental health visit – an appropriate ADHD followup with an inpatient stay seem to me, I couldn't quite figure out why they were using that?

Female: To the best of our knowledge all of our codes are outpatient.

David Keller: OK, hold on. Let me find my comments because I wrote down which one – I went look them up and not a big coding (freak), so I didn't – don't usually do that but it's struck me odd, we find it. Sorry.

Female: Actually and by the way, I just hold up to specification because I had to remind myself, it's a primary or secondary ADHD diagnosis. So it does allow for that.

David Keller: Oh, thank you.

Severa Chavez: OK. That is very clearly stated incorrectly. And so we need to make that very clear because I think you're going to get push back on that and needs to be clear.

Female: OK. So I'm very sorry. I just knew I have to – we went through many (iterations) of testing of the file, what ends up being the final specification. So I needed to refresh my memory. I apologize for not getting to it faster.

David Keller: Yes. So we apologize from this...

(Jill Moore Gordon): So this is. I'm going to say – this is (Jill). I look at the CPT Code 99221 which is a level one inpatient admission.

David Keller: Right, that was the one. I've got a bunch of them listed in my comment. 9921, 221, 212, 223 …

Male: And these are …

(Crosstalk)

Male: Yes.

David Keller: 99233, 99238, 239, 99251 to 99255, all look like they were inpatient codes to me. So...

Severa Chavez: OK. We'll get that list to developer and they'll need to respond at the end of person meeting.

Female: OK.

David Keller: Yes.

Female: Response to that.

Suzanne Theberge: OK, with, no more questions, we can move on to feasibility at this point. The data first for the measure, your e-mail is administrative claim data. For using usability, the measure is not in use.

The developer has not indicated any specific plan for the measure use in public reporting and payment program. The developer stated no unintended consequences. That's an enough...

Female: Let me clarify that. We do not have the measure in use but the PQMP program overall is designed to through the (CHIPRA log), to build measures that can be applied in the Medicaid program.

So if endorsed by NQF, this measure will immediately be a part of the group of measures that's medicaid is asked to evaluate using. So in that, at this time, but it has a pathway.

David Keller: Yes. And I would just say simply state on our state (theme) committee that's in the middle of picking measures. We are hungry for a decent pediatric measure. And we were not happy with the current list of available ADHD measures. So if one this gets approved then we'll be looking at that for all program as well.

Female: Yes. Yes. There are lot of from – at the current list.

Suzanne Theberge: OK. If no more questions for the developer, we would like to move on to the next measure which is measure 2805 pediatric psychosis timely inpatient physiatric consultation.

So the description of the measure is percentage of children or adolescent ages greater or equal to five to less than or equal to 19 years old, admitted to the hospital with psychotic symptoms who had a psychiatric consultant in-person or by telepsychiatry within 24 hours of admission.

The (inaudible) were stated that the evidence supporting this measure the rise primarily from the American Academy of Child and Adolescent Psychiatry 2013 guidelines at Cochrane review and a review of the literature about the developer.

The question we have for the committee as the developer provide us sufficient evidence between the relationship of this measure to patient outcome for the

timeframe 24 hours that must be met to achieve the measures specification for the age range. And if the committee concludes the empirical evidence is not sufficient list specific, does the committee wish to consider the insufficient with exception)path)?

Opening up for discussion.

(Jill Moore Gordon): So this is (Jill). I had huge problems with this only because part of the – it's very – so part of it is related to young children for whom the presentation of psychosis could be a lot of things that are not schizophrenia, and a lot of things who – the addition of child psychiatry to the mix may not be helpful.

I mean I remember seeing a kid who was five who had chicken pox and was on Benadryl and getting Caladryl lotion and she was eating the Caladryl lotion and she came in and she was psychotic. But she had a really high Benadryl level. And once that was gone, it was gone and he was fine.

Male: Right.

Female: Right.

Male: And he – looking for drugs and abuse and that six years old doesn't actually team like the best approach.

David Keller: Well, it's a sort of drug that drugs of abuse?

Male: That's …

(Jill Moore Gordon): That's the next one.

David Keller: That's the different.

Male: That's the next measure. Sorry.

David Keller: Different measure.

Male: Yes.

David Keller: Yes. I think...

Robyn Nishimi: They all – they flow together that they can turn ahead with this one is that, the evidence is coded is about early treatment. Not about – an early treatment being early in the course of disease and this is not about the course of disease.

The first 24 hours at the presentation in the emergency or maybe actually relatively late in the course of disease.

(Jill Moore Gordon): Well, yes or it may be acute and you have to rule out some other cut.

Robyn Nishimi: Yes.

Female: Any other concerns related to evidence?

Suzanne Theberge: I think there (simple).

Male: I have to say that I felt very uncomfortable with the evident. The evident supporting this measure.

Female: Yes, that's what I would just going to say too that I felt that it states that none of the studies actually measures being team with the 24 hours based on a more positive outcome. It wasn't specific.

Male: As I recall also, the evidence sided. It seems within 24 hour in combination with other intervention. If I'm remembering...

Female: Correct.

Male: ... this measure correctly.

Female: Yes.

Female: OK. The developers heard the concerns. Let's move on Nadine.

Nadine Allen: OK. Moving on to gap and care and disparities. The measure captures two elements of performance, access especially psychiatric care for these patients in timeliness of that access. Developer tested the measure using data aggregated over two years from three children's hospitals. The mean performance was 88.5 percent and the range was 77.6 percent to 95.1 percent

with respect to disparities, the developer did not find statistical significant difference in performance across the demographic groups it examined. Is there a gap, so the question for the committee is, is there a gap in care that once a national performance measure. If no disparities information as provided, is the committee aware of evidence that disparities exist in this area of health care and should this measure be indicated as disparity sensitive.

Female: So I thought that was really hard to gauge using this, it was only three hospitals or just children's hospitals, I mean there's a whole at, more out there where the practice maybe different. I mean it's hard to say – and clearly there are disparities around mental health issues. But, you know, I think it was hard to tell if there were disparities in terms of access or other kinds of things.

Nadine Allen: Yes, I though the use of three urban children's hospitals would probably not representative of the typical mental health, initial mental health hospitalization in more rural communities away from major city centers that we know do not have mental health providers especially for children.

Female: Anything else? Yes.

Nadine Allen: What I'm saying is, I do think that there are probably gap that this didn't really bring out.

Female: Yes, in fact they've found, the average was supposed to 90 percent.

Nadine Allen: Yes.

Female: That's not that bad.

Nadine Allen: Right.

David Keller: No. That what's what I was going to say is, I was looking for the room for improvement.

Female: And the Cochrane review – I'm just looking at the Cochrane review to say, hold in it, it doesn't have anything to do with 24 – within 24 hours of presentation. It had to do with treatment during program of schizophrenia.

Female: This is the developer, we would like to respond at all or do you want us to...

David Keller: Yes. So I was going to say, let's turn to the developer and see – my thing as I'm looking at this and say, "This is very performance. It's an opportunity for improvement, or do we think there's variation outside of this – there were kind of the, three hospitals who were able to work with.

Rita Mangione-Smith: Yes, so our respond to that, David. This is Rita Mangione-Smith.

David Keller: Hi, Rita.

Rita Mangione-Smith: Hi, how are you? So the three hospitals where the measure was tested were all as mentioned previously, urban children's hospitals. And the assessment of our center group was that, we still want to put this forward despite the relatively high performance in our field tests because it's our, at least, to believe her – mental health working group that if we had reached out beyond children's hospitals to the rural community settings or even just community hospitals in urban settings or suburban settings. There – we might see very different performance on these measures. We were limited financially from our cheaper grant and how much testing we could do on these measures and you know, therefore it couldn't do a more comprehensive representative tests in various (inaudible) hospitals that have inpatient psych units for children. So, that's why it was put forward despite the high performance we saw on our field tests.

And, Naomi , I don't know if you want to comment on the Cochrane review.

Naomi Bardach: Yes, can you hear me?

Rita Mangione-Smith: Yes.

Naomi Bardach: Great. Hi, this is Naomi Bardach, I'm from UCSF and part of the mental health working group. That Cochrane review, you know, the evidence – there is no specific evidence that support of the 24 hours window. That was something I believe and whether you can tell me (inaudible) about this. But when we put it to our (Delphi panel) that was – actually sending it – people encouraged from the (Delphi panel), actually to look at the 24-hour window.

And there is that dearth of evidence, so I think you guys are seeing the reflection of the evidence summary that there's a dearth of evidence on this particular process of care. There's just not a whole out, not a whole out there but we did feel like it was enough that it would support the development of this measure and – suggesting – and seeing whether or not there is variation across sites.

Nadine Allen: And one other thing, I wanted to clarify earlier was mentioned that what if a child came in because of an ingestion and that's why they were psychotic. The selection of cases for this quality measure come from administrative discharge diagnosis data, so you would have to have a primary discharge diagnosis of psychosis, without any other diagnosis indicating ingestion or a good reason for psychosis. So it would have to be a primary diagnosis of psychosis.

Female: So that makes sense in terms of making sure that people with psychosis that are being evaluated. On the other hand, it seems like the quality issue would be, if you have a little kid in the emergency room, who has psychotic symptoms, do you need more than just the emergency room (keen). Is it appropriate for them to have a psychiatric consult?

Female: And when would it be appropriate they have the psychiatric consult, is it after you ruled out the medical causes?

Female: Yes, maybe that's right. And maybe it's appropriate to just these kids who ultimately have a diagnosis psychosis. On the other hand, you may not be able to rule out the medical causes within 24 hours, it makes (inaudible) 24 hours to get through all of the drugs screening.

Male: And not just drug screening as the center that's there's a lot of an MDA receptor antagonist types of presentation. We've had a number of kids that it's taken quite a few weeks to get those results back.

Female: Yes, yes. So I can step – one of concerns I have was the issue of it just being the discharge diagnosis and I understand that, but the population you're worried about that the kid sitting in front of you in the emergency room and is

it, should you be calling a psychiatrist if you have a kid who's still psychotic and it's been 18 hours, or is there still more medical work up it's appropriate before you need to call in a psychiatrist?

Female: Sure.

Female: I was originally the time window was longer but we've had (Delphi panel) that really pushed us to shorten it to 24 hours for children admitted to the hospital with psychosis.

Male: Can I ask a question about the longer time course because in the measure specification that we got, there is a 24 to 48 hour out as long as somebody justifies documents that there was a reason for delay, the justification for the delay is not necessarily need to be specified. So I see that as sort of questionable or problematic for this. It certainly would allow for perhaps the types of testing where you may need a little bit more time.

But it's also – it's a pretty large out because somebody has to do with, say, I came at 40 hours and it was just to find and that would allow them to be out, or to be included as enumerator. So I wanted some clarification on that.

Nadine Allen: Yes. So, that was something that, in developing this measure, we felt was important to allow for it, that if there were some justification that was documented as to why the child had not been evaluated within the time window that we would not be inappropriately scoring somebody as not having met the measure. It's some kind of a reasonable justification have been provided for why they haven't gone or given the care as proposed to the measure.

Male: All right. So I guess the question is that, at least in that information that we've got, it didn't have what would be a reasonable justification, it just had that, somebody had to have a reason. So is there any thought about that, does that need to be flashed out and made more clearer or is it more clearer in the abstraction piece that's not necessarily clear in the (inaudible)

Female: Yes. I'm sure this – positive examples were given to the abstractors of what – as what would count as a justification. And I can tell you on these field tests,

the finding that was a rare event and that probably is part and parcel to why you're seeing such a high pass rate. By in large that pass rate was people truly getting the consultation not justifying that having gotten it.

Female: OK. I'm going to interject here because we're getting towards out (witching hour) and we still have one measure to go.

Is there anything about measure that the committee wants to discuss amongst itself in terms of reliability and validity? We've heard concerns about the evidence. And we'll set aside using usability for the full committee discussion. So reliability and validity, and I'd like to limit it to the committee right now.

Female: Yes. I did have a concern about the attempt to do an empirical validity measure was pretty challenging because – overall lengths of stay was the only thing that have a very modest relationship in the others had no relationship. And it actually isn't surprising that 38 readmission or 38 return to the emergency department would be change by having – had psych consult early in your first visit.

So I think that there's a little bit of struggle there would being able to come up with something that would be of validity measure. But I'm not sure that I'm satisfied that there is a validity measure yet for this.

Female: Anyone else?

Female: I would add to that.

Robyn Nishimi: Thank you. Let's quickly move to the last measure. Nadine, can you just quickly describe it and then open it up to the committee.

Nadine Allen: We have two more measures, Robyn, sorry. The next measure is 2806, Psychiatric Psychosis Screening for Drug of Abused in the Emergency Department. The evidence is provided using the 2013 guidelines from the American Academy of Child and Adolescent Psychiatry. Developer provides no additional reviews or literature indicating no studies were identified since AACAP published the guideline in 2013.

Any comments from the committee members?

Robyn Nishimi: And what I'd like to do here is if you have questions of the developer, let's raised them but rather than have the developer respond, developer will be prepare to respond at the meeting. But this way, we'll at least get a little bit of the questions aired from the committee.

(Jill Moore Gordon): So this is (Jill), I just wanted to echoed what we said with the last one was that there lots of other drugs that can cause psychosis and while I recognize this as measuring substance abuse in the presence of – or it maybe measuring substance abuse in the presence of schizophrenia. I do think because there are other drugs that sort of overlap that might be a consideration in terms of making sure those were measured.

Male: I wanted to raise the question about the – and/or piece of the numerator and this sort could be some clarification about that.

Robyn Nishimi: OK.

Male: In terms of why is there an indoor none or all are none and I think you had the question and that the other people or somebody else had raise was – I lost my train of thought, sorry. It will come back to me.

Robyn Nishimi: Anything else from the committee on this particular measure?

Male: So if in terms of the first that performance gap – I was –was interested in the extremely lower rate of testing which actually made me wonder whether there was some good reason for a very low rate of testing, and I think perhaps sort of – very – with younger kids, the fact that there are other number of other potential reasons for that presentation to occur, that would be look for and not necessarily looking for drugs of abuse alcohol.

Male: Right. And actually I just remembered my other train of thought, and I think a number of people commented on this summary piece. It's the – actual guidelines that are coded allow for one exposure to drugs of abused cannot

otherwise be ruled out, and it raises the question of what exactly does that mean for the reliability of the measure on the documentation.

Robyn Nishimi: OK. So I think those are all very fair and good questions for the developer to be prepared to answer at the in-person meeting.

Nadine, the last measure, same ground rules just to committee raise in questions for the developer to be prepared at the in-person meeting.

Nadine Allen: So the next measures 2807, Psychiatric Danger (Dissolve), the Discharge Communication with Outpatient Provider. The developer provides evidence that link measure in process of care to reduce in representation with dangers of suicidality.

Is the relationship of this process measure to patient outcomes clear? Given the information submitted, how should the developer literature review be assessed as a systematic review or as a empirical evidence submitted, or is insufficient with exceptions rate in appropriate.

Robyn Nishimi: Nadine, I'm sorry to cut you off but I want the committee to also address whether they have concerns about the reliability of the specifications and the testing.

Female: So I had a strong concern about the quality of the evidence because it's – there are systematic reviews but a systematic review of not very good evidence, still remains not very good evidence.

And I just didn't think was very high at all to support a measure. To support the idea that there ought to be some communication, I don't think that's one of those things that doesn't need evidence. But whether it ought to occur between 24 and 48 hours before the discharge and whether it ought to be done in the particular way. I just don't think we have enough information.

David Keller: (Boy). So this is one of those ones where I agree with you that there's not a lot of evidence, but having been the guy out in a little town in Central Massachusetts who has the patient walk into his office saying, "They put me on five meds

and I'm running out, I need refills." For a patient, I have no notion with psychiatrically hospitalized.

I can tell you that happens a lot.

Robyn Nishimi: Oh, I know. I've been in practice, and I've had this experience but it doesn't mean there communication between 24 and 48 hours prior to discharge, which is the specification is going solve that problem.

David Keller: Well, it would make it a lot of easier if I knew the patient was coming out of the hospital. I mean, they handle a piece of paper and say that they, you know, take this to your doctor, is not good.

So I saw this despite the lack of evidence is something that, boy, I would love this to be a measure because it would take a huge. It would push the psychiatric hospitals into doing something other than saying that they have the current practice which is to say what we made communication available. Did you scream into the woods and nobody's listening? That's not communication.

And so I actually – we can discuss – we'll discuss this I'm sure at the bigger meeting. But I think that's the – that one doesn't throw – that part of this, this measure doesn't trouble me quite so much. There were something else I was trying to find my notes, there is something else in here that did.

Female: So I think it's very likely that this would be missing from the chart. I think it's a likely. So in terms of the reliability is I think that the communication could have occurred and would not have been documented and under the feasibility, very, very difficult to extract this data.

Male: Then I had – I concerns also about in terms of the obstruction, who would be responsible for forwarding the information and what information actually it would be forwarded or relate. Is it the case manager? Is it the social worker? Is it the position for clinician extender? Is it the physician if you have teaching – if you have teaching – if you're a teaching facility? Is it the resident? And who is it being forwarded to? And if it's forwarded to a

primary care position but not necessarily, you know, there's all kind of questions around that I really wanted to see form the abstraction tool.

Nadine Allen: And I think the other point about, who is it being forwarded – to? David, I'm not sure you would have gotten it as the primary care if there was a behavioral health person involved. So it would be helpful to know who that supposed to go to.

But another concern I had about it was – the exclusion of practitioners that were in the same system and that sort of assumption that you're going to get it through the medical record which I think is not a good assumption

David Keller: Right. I was – that was the part that I didn't like actually. I thought this should include because I don't think communication within organizations works any better than communication via outside.

(Crosstalk)

Male: And those are the expression of the patient, this (AMA), I thought – I know that some of the summary comments were of kind of the same. You would definitely want to make sure that information got to somebody who is likely to see that person as an outpatient.

David Keller: When the surface, I guess – yes.

Female: Well…

Male: I had a question about – there was an inconsistency that I think I read that it's the metric. The measure is to discuss the plan or to forward that information prior to this chart. But it allows for – the numerator allowed for to be done after 48 hours after discharge. But it allows for – the numerator allows (inaudible) to be done up to 48 hours after discharge, so there some discrepancies and the description versus the numerators specifications.

Suzanne Theberge: And the other thing I think about it that one of the measures that they often use around discharge from hospitals specially psychiatric hospitals is, actually, a physical followup post discharge. And why you can't measure that

from the hospital discharge record or from the hospitalization record is, you know, is that a better measure than just the communication?

Robyn Nishimi: OK, all terrific comment. Suzanne, I think we only have two minutes, so I'm going to turn it back to you. And I apologize to the committee to kind of short-circuiting your discussion on these last two measures. But I think the developer has a feel for what they need to address at the in-person meeting.

Male: Because if we just make…

Female: Sorry, this is the developer, actually, could you do as a big favor and – there were (fast and furious) comment, all very much (appreciated). But I don't – I'm not sure, I can be confident answering…

(Crosstalk)

Female: OK, awesome. Thank you.

Suzanne Theberge: We'll have a recording and as well as a transcript.

Female: OK.

Female: OK.

Suzanne Theberge: So on (inaudible) very brief public comment and then I'll just very quickly run do the next steps. Operator, can you open the lines for public comments please?

Operator: Certainly. In order to ask or make a public comment at this time, please press star one.

Suzanne Theberge: And you can also submit comments via the chart box if you're not on the phone.

Operator: There are no public comments at this time.

Suzanne Theberge: All right, next slide please. Very quickly, we'll just go to the next few steps.

So thank you so much for all the time and efforts you put into reviewing the measures today. It was great discussion. We'll be doing similar discussion at the in-person meeting. So at this time, we'd ask you to start reviewing your remaining measures in the project.

We will be assigning you, a few people to each measure. I believe discussing meeting you'll be responsible for kicking off the discussion and being very familiar with the content to the measures. And we'll be getting in touch with you this week about what measure you will be – that (inaudible) discussion on.

You're travel arrangement should be all set. And you should have received – I think we should be sending a hotel confirmation and information today or in the next day or two. So please, let us know if we have any concerns about that. Next steps – I'm sorry next slides.

We have one more work groups call on Wednesday you are welcome to listen in. But you don't have to attend, unless you would like to. And then, we'll see you all right after Thanksgiving at our office in Washington D.C. December 1st and 2nd.

And then, if we don't get through everything at the meeting, there will be a follow call in December 10th to finish discussing the measures. Next slide.

This is our contact information. Don't hesitate to give us a call or e-mail at any time, if you have any question. All measure information will be posted on the SharePoint side for the committee and we'll be updating those measures worksheets that you've been looking at. In the next couple days, we'll be getting and all the comments from the work groups so that you can see what your colleague have been thinking about the measures. And so we're at time, so I will stop there and just quickly see if there any question?

Female: Really terrific discussion today, we will really appreciate it.

Male: And thank you everyone. I hope – thank you for putting this together and thanks to our measure developer for hanging with us particularly Rita who hang in for a longtime 'til she can...

Rita Mangione-Smith: Yes, thank you.

Female: Thank you, everyone.

Female: Bye-bye.

Suzanne Theberge: All right, thanks everybody.

Female: Bye-bye.

END