NATIONAL QUALITY FORUM

Moderator: Performance Measures Pediatric
November 18, 2015
2:00 p.m. ET

OPERATOR: This is Conference ID #16878526

Welcome everyone. The webcast is about to begin. Please note today's call is being recorded. Please stand by.

Suzanne Theberge: Good afternoon everyone and welcome to the Pediatrics Performance Measures Committee Workgroup Call Number Four. Thank you so much for joining us today.

Before we begin, I am going to do some housekeeping announcements, and then we'll do introductions, and then we'll begin. So, first and most important is that committee members and developers, you must be dialed into the phone line to speak. We can't take audio input over the computer. So please do dial into the phone number on the agenda. And if you were dialed in and streaming the webinar on your computer, please turn the volume off on your computer because the audio will stream and the interference can be pretty confusing.

We do request that you put your phone on mute if you're not actually speaking to reduce feedback, but please don't put us on hold. This will get hold music on the line.

And, next slide, please. I think we'll just dive right into the introductions. As I said, this is Suzanne Theberge, I'm the Senior Project Manager on the Pediatrics team. And I'd like to ask my colleagues at NQF to introduce themselves. Nadine?

Nadine Allen: Hi, I'm Nadine Allen. I'm the Project Manager for this project.

Robyn Nishimi: And Robyn Nishimi …

Suzanne Theberge: Robyn?

Robyn Nishimi: Robyn Nishimi, I'm a senior consultant to NQF.

Suzanne Theberge: All right, thank you. And we also have Severa Chavez on our team but she is not in today.

So, next slide. I would like to go through the workgroup member introduction. I'm going to just call roll and ask you to introduce yourself in just really briefly, a sentence or so after I call your names.

John Brookey?

John Brookey: Hi, it's John Brookey. I'm from Kaiser Permanente. I'm a pediatrician and a quality medical director for Southern California Kaiser.

Suzanne Theberge: Thank you. James Duncan?

James Duncan: Hi, this is Jim Duncan. I am vice chair for Quality and Safety Department of Radiology. And I'm a practicing interventional radiologist spending my time between St. Louis Children's Hospital in our adult facility. And I have to say that I have to recuse myself from the discussion on Measure 2820, the pediatric CT dose.

Suzanne Theberge: Thank you. Deborah Fattori?

Debora Fattori: Hi, this is Debbie. I'm a nurse practitioner. I'm still practicing in the Department of Kidney Transplant. I am the senior director of Ambulatory and Advanced Practice Nursing at the DuPont Hospital for Children, and I'm also representing the NAPNAP organization, the National Organization of Pediatric Nurse Practitioners.

Suzanne Theberge: Thank you. Ricardo Quinonez?

Ricardo Quinonez: Hi, I'm Ricardo Quinonez. I am a pediatrician with the Baylor College of Medicine, and a pediatric hospitalist and chair of Quality at the Children's Hospital of San Antonio.

Suzanne Theberge: Thank you. Jeff Schiff.

Jeff Schiff: Hi, this is Jeff Schiff. I am the medical director for the Minnesota Medicaid Program and a pediatric E.R. doc practicing at Hospital of Minnesota in St. Paul.

Suzanne Theberge: Thank you. Jeff Susman? Jeff Susman, are you on the line?

All right. And Keith White?

Keith White: I'm Keith White. I'm a pediatric radiologist. And I'm medical director of imaging services for Intermountain Healthcare.

Suzanne Theberge: Great. Thanks, everybody, and so much for joining us, and we do have a number of measure developers on the line today. (Most of these) measures are under review. So I'll ask you just to introduce yourself as we put your measures if we have questions. And as folks are talking, if you could just say who you are as we all get to know each other, just say your name before you start talking, that would be great. Thank you.

All right, next slide. So, the process for today's call is that we are going to go through the four measures that we're looking at today. We're going to walk through the criteria for each measure. I will be kind of leading the discussion in that I will be flagging issues that you all raised on your pre-meeting surveys, raising questions that NQF staff had noticed in their preliminary review of the measures, and asking you all to kind of expand on the thoughts that you have put into the survey. We will not be voting on any of the criteria today, this is just discussion.

And because we are not voting, we will go through all the criteria whereas at the in-person meeting, if you are discussing a measure and you vote and it does not pass, it will not go forward. At this time, we will discuss all of the criteria.

We are really going to focus our discussion on either places where there were big questions or where there were different opinions that were among the various committee members. We're not going to spend a lot of time on places where everyone agreed. If everyone agreed that a measure is generally feasible, we'll just kind of skip past that.

Again, the developers are on the line. If you have specific questions, please do request that the developers answer them. And – but that said, this is the time for the committee to discuss, so we're not going to be having the developers do any kind of introduction of their measures. We'll save that for the in-person (meeting).

So, after we go through all the measure discussion, we're going to allocate about 25, 30 minutes per measure for today's discussion. We'll do a brief member and public comment if anybody's on the line and wishes to make a comment, and then we'll go over the next steps.

All right. So, next slide, we have one mental health measure on today's call and then we've got three radiology measures. And we're going to start our discussion with measure 2815, that is mental health follow-up timeliness, delayed coordination of care. This is a measure from the CAPQuaM Center of Excellence and looks at the presence or absence of delay in follow-up visits with mental health and primary care clinicians following hospital discharge of a child with a primary mental health diagnosis or from a mental health facility.

So, we did see some concerns raised here around the evidence of the measure. And I want to see if anybody – I'll pause here and see if anybody has any questions – actually, I'm sorry, I should have done that first. Are there any questions before we start discussing the measures?

Jeffrey Susman: This is Jeff Susman. I just joined.

Suzanne Theberge: Oh, great, thanks, Jeff. Could you give us one-sentence introduction?

Jeffrey Susman: Sure. I'm the dean here at Northeast Ohio Medical University and have a strong interest in the interface between mental health care and primary care.

Suzanne Theberge: Great, thank you. All right. So, we'll go dive right in to 2815. So we did have some questions and concerns we (used) around the evidence on this measure. As you know, the first criteria for the measures is importance to measure and report, and we look at two sub-criteria -- one being evidence and one being gap.

Does anybody have anything they wish to discuss about this? Anyone want to start the discussion off?

Jeffrey Susman: Well, I mean – this is Jeff. There is not clear evidence linking the outcome to the measure. And while it might have some face validity or themes like it should at least in the documentation presented, there really wasn't a good job of providing that in my mind.

John Brookey: Yes, this is John Brookey. And I don't know if the measure developer is on the line or how we're going to do this, or we're going to be asking the question …

Larry Kleinman: I am on the line, just to let you know.

John Brookey: OK. So I read through this a couple times, and again, I'm very interested in this topic. I think it really makes a lot of sense clinically, kind of intuitively, but in terms of any hard core evidence especially related to the primary care follow-up, I mean, I can kind of really kind of understand more the behavioral health follow-up. But, in our population, I don't know that it necessarily, there's going to be a strong evidence for also that requirement for the primary care follow up. And I just didn't know if there's anything out there that would really support that or is that just something that we feel clinically (inaudible).

Larry Kleinman: So this is Larry Kleinman. I'm sorry, that was a question, yes?

John Brookey: Is that OK to ask questions now or should we just be making comments?

Suzanne Theberge: Well, we'd like to primarily have you make comments. But if you do have a specific question, then, yes, do ask that question and we can have the developer respond.

Jeff Susman: I mean part of the …

(Crosstalk)

Jeff Susman: … has to do with the evidence around timing of these things. And, again, I think while one might makes some sense intuitively, it seems like this was pretty much pluck from the air.

John Brookey: Is that Jeff speaking? Who was speaking?

Jeff Susman: Yes, this is Jeff Susman again.

John Brookey: OK. Yes, that was my opinion as well.

Deborah Fattori: So this is Debbie Fattori speaking. Certainly, I agree with the comments that were made so far. It does seem to make clinical sense. I've been spending some time out by institution working on readmissions in tertiary and (cautionary) pediatric health care readmissions not necessarily mental health, and certainly transition of care, and follow-up plays an important role with that. But I'm just not clear that the evidence presented was enough to support the potential outcome measure or the outcome of avoidance of follow-up issues and harm to the patient.

Suzanne Theberge: All right. Are there other thoughts about evidence for the measure? What do folks think about the gap? Is there a gap in here that requires a national performance measure?

Jeffrey Susman: Well, there's certainly a gap in outcomes of the – and there's a variation in what is happening. But is that really a gap in care that is meaningful or not, I think, isn't well delineated. I mean I would have preferred a much crisper causal pathway to specific outcomes of interest to patients or health care system.

Suzanne Theberge: OK. All right.

Larry Kleinman: Is this a good time to respond or – it's Larry.

Suzanne Theberge: Well, I am kind of hearing general comments, not specific questions. So, unless the committee member has a very specific question, then, no, we'd ask you to kind of hold on that.

Larry Kleinman: OK.

Suzanne Theberge: So, all right, does anybody have any further comments about evidence or gap, or shall we move on to the testing of the measure, the reliability and the validity?

Jeff Schiff: This is Jeff Schiff. I just need to ask – I just want to ask a question about why one visit within 30 days was chosen because that's, you know, sort of – I guess the question is, were there other looks at multiple visits or a shorter time frame that might have been associated with readmission or outcomes for patients.

Larry Kleinman: OK to respond to that?

Suzanne Theberge: Yes, Larry, please do.

Larry Kleinman: OK. This is Larry Kleinman, I'm the P.I. of CAPQuaM and a general pediatrician. This – when we develop the measure and the measure we submitted to the Pediatric Quality Measures Program rather than to NQF, we had following sources of evidence prior to the development of the measure. We had lit reviews which described continuity of care as care within the mental health system; coordination of care, which the literature defines as a provision of care across specialties and/or sites and other things; and, we did interviews with clinicians and we did focus groups with families, or a focus group with family. So that was the evidence that was then summarized and presented to a carefully chosen interdisciplinary national expert panel.

In the panel, there were several timeframes that were identified as important. Timely care was defined as mental health follow-up within 21-day, sorry, within seven days and primary care follow up within 21 days; so seven and 21. Delayed care was defined by the expert panel using a RAND process as greater than 30 days. And then, there was failure which was defined, I believe, as 90 days. I don't have that right in front of me, but these all were

looked at, they were all analyzed. And when we were discussing with the folks at NQF and AHRQ about what to submit, it was clear we were only allowed, going to be allowed to have one measure on this agenda and we weren't able to submit the various matters.

And in speaking with our clinical experts, this delay measure was felt to be most critical. We did find that for primary care, that delays in primary care were associated with higher rates of readmissions. We didn't develop sophisticated models, we could do it. Although I don't know that we can do it in the time frames in this because we don't own the data. The data is all handled and analyzed at New York State Medicaid. But – so we did some cross tabs on things.

Now, interestingly, you do not find that same – if you're looking at measure's essential tendency, you do not find the same relationship of readmission associated with mental health follow-up, but of course, that's to be expected because the most sick children, the ones who are at greatest risk for readmission are also the most likely to have very timely. (Good thing) I would say in terms of gap is this the request to develop this measure came from AHRQ and CMS. So the topic was there as the measure hours. And it was an attempt to take the current NCQA measure and to optimize it for pediatric population. As you may know, it incorporates some of pediatric populations, not all but, you know, it is the same identical for mental – for adults and for children.

So, the numbers, we analyze that different cut points, the judgments were made by an expert panel, we considered scores of eight or nine rather than seven, eight, or nine to the strong endorsements of it and then we also – we adjusted for disagreement and looked to make sure that it was still at least the seven for anything that we would use. We gave them various opportunities. It's not that we said, "Is 30 days a good time?". We said – we gave them, 15 days, 30 days, 45 days, 60 days, et cetera. We gave them a range and ask them to rate all of them independently.

So, this came out of a very formal process with a formidable group that we had pediatricians, psychiatrists, child psychiatrists, psychologists, family

physicians. I think we even had an informatician because we thought that part of follow-up in developing measure is related to follow-up and this is not – you know, timely, this is not the only area we'd considered, but it's what we (Audio Gap) that informatics might be an important part of that. Part of what we wanted to make sure was that anything we ask was something that the discharging hospital and/or the system had an opportunity to impact that these were things that were modifiable and changeable as well as clinically meaningful.

So there was a very systematic approach to it. And the primary care which the panel felt was very important also did hold up in terms of having relationships with readmission rates as you extended out to the longer end of the spectrum. That was less true for the timely or not. But as you get to the issues of delay and failure, we saw that. Hope that's helpful.

Suzanne Theberge: OK, thanks, Larry. Yes, thanks, Larry. Are there any further questions? All right …

Jeffrey Susman: It would be helpful – I mean, to be honest, not having that documentation without understanding that evidence, it's really hard to judge the evidence, I mean, the summary that was provided. And while I can imagine the sort of rigorous process that you've used, my concern is frankly we didn't have that.

Larry Kleinman: How best can I present it? Should I share that in the in-person meeting? Should I summarize it or have a handout …

Suzanne Theberge: Let's see if I can – why don't we follow up with you after the call, Larry? We can …

(Crosstalk)

Larry Kleinman: You got it. Thank you.

John Brookey: So this is John Brookey. Yes, I had to – I was distracted for one second, but what it sounds like you have more of a consensus-based approach to this rather than evidence based. Is that fair to say?

Larry Kleinman: CAPQuaM is not like other measure developers. In that we developed a group – a high-level group of stakeholders including professional societies and NCQA and others, New York Medicaid, and developed a consensus around the process. The process did not start with a measure and asked the group to validate the measure. The process started with a topic and asked the – for the critical constructs within that topic area to come out of this modified (Delphi) method, which I hear wasn't a true consensus procedure but a consensus-related procedure.

Jeffrey Susman: So you're using the RAND Appropriateness approach to …

Larry Kleinman: We used a slight modification. We added some additional information in because we did those interviews and focus groups. And we also asked them to rate the constructs where things were no longer mutually exclusive. So, 30 days didn't mean that 15 days was no good. If 15 days was timely, then seven days was also going to be timely. If …

Male: Yes.

Suzanne Theberge: OK, Larry …

(Crosstalk)

Suzanne Theberge: … we got you cut off because we've only got 10 more minutes to discuss …

John Brookey: Yes.

Suzanne Theberge: … your measure.

Larry Kleinman: OK, sure.

(Crosstalk)

John Brookey: And I'm not trying to be – I'm not being critical. I just think that if that's what it is, then it just needs to be stated and the process needs to be described. Now, it's an important issue, it's just that I think that we're trying to struggle between is there evidence or is it more of a consensus based.

Larry Kleinman: Right.

John Brookey: And if that's what it is, just describe it for us.

Larry Kleinman: Yes, I'm happy – we were asked to cut down what we presented and maybe we didn't do that well, maybe we didn't make the right judgments on where to cut, so. So, I'm happy …

(Crosstalk)

Suzanne Theberge: OK, we'll follow up with you after the call. We can discuss providing further materials prior to the in-person meeting for the committee to review. But at this time, given our, you know, our limited time for each of these measures, I'd like to move on to the reliability (Audio Gap). And there were some questions raised here as well. Would anybody like to kick off the discussion or shall I summarize what was submitted in the comment?

Well, I did hear – I did see some concerns raised about the reliability testing. Folks were not sure that empirical reliability testing was conducted or validity testing there. They have some questions there as well as with the use of claims data. Anybody want to be more specific about your comments?

Jeffrey Susman: The actual – again, the actual testing used, for example, for reliability testing, is not really described and there really no actual results. It's not that I doubt that this was done, there's too much sketching is for my way of thinking to accurately judge whether this is an issue or not, for example.

Suzanne Theberge: I saw some concerns about the use of administrative claims data, whether that accurately reflects mental health diagnosis, diagnosed disease, and self-injury, suicide attempts or suicide idealization. Sorry, you all know …

Jeffrey Susman: Ideation.

Suzanne Theberge: Thank you. Tongue-tied thing.

Jeffrey Susman: It's all right.

Suzanne Theberge: Do folks want to speak about that further? I saw that comment come up several times in our …

John Brookey: You know, it's interesting – this is John. We're creating a suicide safety net here at Kaiser, Southern California. And I just pulled coded data last week. Thinking of it, (I find) nothing, and I actually found several thousand codes for ideation in various kinds – there's about 75 codes or something that go along with this. So, I don't know, I think it's probably undercoded but I did not find a lack of codes. I found plenty of people that are coding this diagnosis at least in Kaiser.

Suzanne Theberge: OK.

Jeffrey Susman: Yes, I think it just depends …

John Brookey: That …

Jeffrey Susman: That certainly would be a place where variation would occur, I would think.

John Brookey: Yes. And what I'm saying is very anecdotal. But there's something out there. It probably is not complete.

Suzanne Theberge: One of the other concerns I saw raised a couple of times was around the exclusion of day of discharge follow up to prevent (gaining). And there was a comment that might punish a best practice or penalize highly performing system. Did that commenter wish to speak to that?

Jeffrey Susman: I've said that, this is Jeff again. And it just seems that maybe the ultimate, the (Cadillac) version is to, you know, discharge the inpatient, have them go over the outpatient clinic, discuss with the hospitalization was about, and schedule ongoing follow-up. You know, I think there's at least some good reason why promoting that connection and continuity, that link back could be a best practice. So, I was concerned that that was excluded, although I understand that, you know, any of these measures probably can be gained by someone who's crafty enough.

Larry Kleinman: Would it be helpful to explain this whole rationale behind that?

Suzanne Theberge: Is that something the committee wishes to hear now or do you feel satisfied?

Jeffrey Susman: I mean, if we had the time, I mean, I think the reality is we're probably already past.

Suzanne Theberge: We've got about five more minutes by my clock for this measure.

Female: Why don't we address it at the in-person meeting?

Male: OK.

Suzanne Theberge: Sounds good.

Male: Will do.

Suzanne Theberge: We'll have more time then. All right. So, are there any other comments from the committee about the scientific acceptability of the measure, the reliability, the validity issues of missing data? It was noted that it was not – did not seem to be included in the documentation.

John Brookey: Personally, I think that it probably can't be measured fairly reliably. That's just my speculation. But, you know, it's hard to tell from this documentation whether that's true or not.

Suzanne Theberge: OK. So, if there are no other comments, we can move on to feasibility. We did have a couple of comments here and a feasibility looks at a, you know, or whether there's a burden on this measure, whether it's fairly available, the data for the measure. And there was a concern that the data would be relatively available using administrative claims but it might be limited at the facility level. Any further thoughts on feasibility?

All right. Hearing none, I think we can move on to usability and use, which is how people can use a measure and whether it's in use and how it could be used. We did see some concerns here. As mentioned earlier, I think this is a new measure so there's some limited information there.

There was a question raised about whether, you know, this could enhance follow up but it is an important aspect of continuity. You know, does the commenter wish to expand on those comments?

Jeffrey Susman: Yes, I mean, just because someone shows up in primary care, if there hasn't been adequate communication, how much value is it. So if I'm in my family medicine office and seeing a patient who discharged for their bipolar hospitalization, but I don't know anything about their meds, I don't know what the follow-up directions from the treating providers were, and don't have any general sense of, you know, suicidality, the course of treatment, blah, blah, blah, then I sort of say, "So what?"

Suzanne Theberge: All right, thank you. Are there any further comments on this measure?

James Duncan: This is Jim Duncan. The related or competing measure, is that something to discuss at this point?

Suzanne Theberge: No, we'll go over this a bit more at the in-person meeting, but I don't think that's actually going to be an issue for this project. The staff might highlight what the related and competing measures are, but when we have two measures that are competing in the same projects, we would ask you to decide on the best in class, but if a measure is not – if you are not looking at another measure, if it's in a different project, obviously, you can't compare the two if you haven't looked at the other one. So, that's not something that I think – well, we won't be discussing it on the call and we probably won't discuss much at the in-person meeting, but good question.

All right. Sorry, go ahead.

Jeffrey Susman: You know, the only other thing that I was wondering about, you know, I didn't spend a lot of time is just how much was this going to yield a – if you (Audio Gap) of conditions, all of which might at least in concept weren't follow-up but may have different levels of urgency.

Suzanne Theberge: All right. Well, I think that we can probably finish up the discussion on this measure. Hearing no other comments, then we'll move on to the next measure. We're going to completely change gears here.

Larry Kleinman: Let me just say thank you all very much.

Jeffrey Susman: Hey, thank you.

Suzanne Theberge: All right, thanks, Larry.

(Crosstalk)

Jeffrey Susman: Sorry to be so difficult.

Larry Kleinman: That's OK, we'll (duke it out) in person, it's fine.

Jeffrey Susman: OK.

Larry Kleinman: So, we'll try to meet the challenges, thank you.

Jeffrey Susman: OK, hey, thank you.

Suzanne Theberge: All right. So, the next up, we have two measures from Q-METRIC, The University of Michigan. Do we have the developers on the line?

Male: Yes, Q-METRIC is here.

Suzanne Theberge: Great, thank you. All right, so the first measure that we're going to look at is 2797, transcranial Doppler ultrasonography screening among children with sickle cell anemia. And this measure is – all right, I'm just – is this a process measure? And looking at the percentage of children ages two through 15 years old with anemia who received at least one transcranial Doppler screening within a year.

So, we're just going to dive right in again and start looking at the importance criteria, the evidence in the – starting with the evidence. Folks had some concerns here. We got a number of comments around the evidence for this measure. Does anybody want to start off the discussion?

We did see folks saying that, you know, there is some clinical guidance behind this, but there was some concern about the age recommendation. The practice guideline, I guess, is ages two through 1, and the measure

specification is ages two through 15. And we did have a number of comments about that and questions flagged

So, do you have any thoughts or shall we turn that over to the developer for a response?

John Brookey: I interpreted that to mean (inaudible) age of 15. So I didn't see the conflict, that means the developer can mention whether there's a comment here or not.

Jeffrey Susman: Yes, it just wasn't as clear, I would agree. But …

Deborah Fattori: This is Debbie. I would agree, too. I mean, overall, I think this is one of the stronger measures that we reviewed in this group, but age was a little bit confusing.

Suzanne Theberge: All right, could we ask the Q-METRIC team just to very briefly speak to the age questions.

(Sarah Reece): Hi, thank you. This is (Sarah Reece) from the Q-METRIC team. We're happy to clarify the ages. The guidelines state that TCD Screening should be received annually from age two and continuing until 16. And our measure is specified from age two all the way through the very end of their 15th year; so all the way up until the day of age 16.

(Off-mike)

Male: Got it.

(Sarah Reece): So there's no conflict.

Male: There's no discrepancy.

(Crosstalk)

(Sarah Reece): Exactly. Thank you.

Deborah Fattori: OK. That's helpful.

Suzanne Theberge: OK.

(Crosstalk)

Suzanne Theberge: Go ahead.

Jeffrey Susman: You want to talk a little bit more about evidence?

Suzanne Theberge: Sure, yes.

Jeffrey Susman: So, why annual? I mean – so if you had five of these annual screenings in a row, is there an appreciable list? Or another way to ask this, what is the natural history of this disorder or problem that occurs? I mean it's clearly, there is a substantial increased risk, so I guess I'm just trying to have a sense of, is there are some diminishing returns? Does the risk change with age? Why not biannual or why not every five years? Presumably, there are some evidence bases that came up with this, but I didn't see that unless I missed it.

Suzanne Theberge: OK, Q-METRIC, can your team answer that?

(Sarah Reece): Sure, hi. This is (Sarah Reece) again. The NHLBI guidelines are very specific, and that children from two to 16 should receive a screen every single year. And there's no indication that that should be stopped or it changed based on the results of previous screening.

Jeffrey Susman: So I guess my question is what's the evidence underlying that NHLB guideline. I mean there is a lot of things that people think but may not be based on high-level evidence, and that's what I'm trying to get to.

(Sarah Reece): There's no evidence to suggest that the risk does diminish. It does shortly after 16, which is why I believe the guidelines go all the way to 16.

Jeffrey Susman: So, to be, again a little more pointed, does that mean that your risk after five negative screens and a stable measurement could appreciably increase in the sixth time? I mean, I confess I'm totally ignorant about this area other than the fact that it's recommended to do. But I don't – again, it's a natural history question.

James Duncan: So isn't it linked to the transfusion therapy and using that as a guide? And that's – you know, there is a screen and then there's an intervention. I'm sorry, this is Jim Duncan, but – and I thought there was very strong evidence in the two randomized control clinical trials and the 50 observational studies, you know, 11,000 patients that this is accepted within this population.

Jeffrey Susman: I don't disagree. It's just the frequency of screening that I'm questioning is …

James Duncan: Right. And you could make the parallel that it's like screening mammography.

Jeffrey Susman: Right.

James Duncan: It's a very hot topic, but I think you have to start somewhere. And annual is, you know, probably – you know, I'm sure they didn't run nine months versus 12 versus 15, you know, to try and assess, you know, which is going to be the best timing. But, that's what sort of comes out later on in the life of a measure like this, I would think.

John Brookey: Yes, clinically, our sickle cell program which has fairly large number of kids, they do do annual screening on the high risk (assessed disease). The one thing that our guys were mentioning, though, is this is talking more about transfusion. There – I didn't see anything in this write-up about hydroxyurea and didn't know if that was just not addressed or why that wasn't mentioned. Probably not that important but I didn't get any disagreement from our experts about annual screening from two to 16.

Male: Yes.

(Sarah Reece): Would you like me to address that point to the developer?

John Brookey: If it's relevant, otherwise, we can probably move on.

(Off-mike)

(Sarah Reece): Sure. That's not relevant to this measure. A clinical trial should the hydroxyurea was not anymore effective than blood transfusions in reducing stroke risk.

John Brookey: Great, thank you.

Suzanne Theberge: OK, are there any further questions or comments on the evidence?

Ricardo Quinonez: This is Ricardo Quinonez. My only question and I think I put that in my comments was, does this apply to every single sub-type of sickle cell disease. I thought I saw some evidence that only certain sub-types are at increased risk for stroke.

(Sarah Reece): That's a great question. The guidelines are specific for children with sickle cell anemia or hemoglobin SS, and our measure is specific to children with sickle cell anemia. In fact, we did actually do some additional work in which we validated our case definition of children with sickle cell anemia against newborn screening rates and showed that our definition was highly sensitive and specific.

Ricardo Quinonez: Great.

Suzanne Theberge: OK. I actually didn't see any concerns about gap here. There was (inaudible) agreement that there's a gap in here. So, does anybody have any comments on that, or shall we move on to the reliability and validity?

OK. So, moving on to the testing, reliability and validity, we did get some questions which I thin may have (Audio Gap). We had a couple of comments on making sure that the diagnosis is correct and – or how reliable is the diagnosis and it sounds like that concern has been addressed. Is that correct or are there any further questions?

Female: That's correct.

Suzanne Theberge: All right. And there were also some questions raised here about the age discrepancy between the measure and the guideline which I think we've also clarified. One question I did see (related to) the comment was saying that validity could be improved by assessing results for why screening was not performed in populations and settings that could otherwise show high level of performance. Any thoughts on that from the committee?

James Duncan: This is Jim Duncan. That was my question and I think it's great that they had data from state to state showing the rise. But I start to wonder if at some point it could hit a plateau that 100 percent might not be achievable in a real-world setting. And that there are, you know, clearly might be some reasons why a family might say, they don't want their child screened, maybe they have an objection to blood transfusions, or there are some other factor going on that would, I guess, penalize an institution that they're not going to get the 100 percent.

Jeffrey Susman: On the other hand, every institution would be under those same challenges, if you would. So I'm not sure it would threaten the validity of the measure particularly, unless you're – I guess, if all the Seventh-day Adventist or whoever that don't want transfusion.

John Brookey: So this is John. I was very impressed with the Medicaid data. I didn't see anything more current than 2010. But I wonder if we know what it is today. And also do we have any data from commercial plans that would compare? You don't have to answer it if you don't have the answer. I'm just thinking out loud about I like to know more about other kinds of payers.

Suzanne Theberge: Yes. Can you or team address that now? Or do you want to bring that information back to this maybe later?

(Sarah Reece): I can …

John Brookey: It was a short – go ahead.

(Sarah Reece): Sure, I can address that just a little bit. We actually do have data but not too much more updated, just 2011, from Michigan that show rates really (Audio Gap) very comparable to what we showed in 2010. We are pursuing looking at this in some commercial plans, but I'm not sure that we'll have that information in time for the December in-person meeting.

And also, most kids of sickle cell are really in Medicaid, probably upwards of 70 percent. Some of that even shows a little higher. So, would you feel this is pretty represented of a very large proportion of these children?

John Brookey: I think it was very compelling data. I just thought it was just great. We really feel like there's real value and, you know, reporting of this kind of information.

Male: Exactly. I agree.

Suzanne Theberge: OK. Are there further comments or questions about the testing of this measure, about the reliability or the validity?

Jeffrey Susman: Just in the reliability. How reliable is the measurement of the velocity itself. In other words, is there inherent measurement viability that would contribute to the overall measure, if you will, being less reliable? I guess this could be briefly answered by the developer.

(Sarah Reece): Sure. That's a great question. The test is generally pretty reliable in terms of the mean blood velocity. However, we weren't assessing any outcomes of the screening just if there was presence of or no presence of a screening as recommended by NHLBI.

John Brookey: So just an anecdotal comment from Kaiser. Our sickle cell program only trusts the results from one of our academic centers in Southern California. I wouldn't want to publicize that, but I think if they believe there is variation. So, that's probably a different question altogether, but I think there is probably …

(Crosstalk)

Jeffrey Susman: Well, it has to do more within setting a cut point. And if the underlying measurement is unreliable of the velocity, then one might suspect you're either over treating or under treating kids at risk.

John Brookey: Correct.

Jeffrey Susman: Anyway, minor issue.

John Brookey: We have a pediatric cardiologist on the call, right?

Suzanne Theberge: We have two. Although, we may have lost one of them temporarily. Jim and Keith, are you there?

James Duncan: Yes. So this is Jim Duncan. Clearly, there's going to be, you know, sensitivity and specificity in the area under the curve that's probably going to vary from lab to lab. Ultrasound, unfortunately, is always difficult to try in standard rigorously. But I think one advantage of endorsing a measure such as this is I think it would, you know, help drive more reliability within the measure. And most ultrasound labs undergo accreditation. It's clearly been done in the carotid where they received that – I mean, there's sort of a halo that forms around measures like this that, you know, most sites would find their good labs that they believe. And it really the treatment is transfusion. And so, I'm not sure if it, you know, were chemotherapy or something with a, you know, an incredible amount of risk, I think it would be a little bit different equation for me.

Jeffrey Susman: Yes. Yes, that's a good point.

Suzanne Theberge: All right. I think we can move on to feasibility. I didn't see too many concerns raised here. The committee generally felt that this measure relies on data that's available, required elements that are routinely coded. Does anybody have any questions or concerns? Or shall we just move right on the usability?

Female: Yes. Me, too.

Male: Sounds good.

Female: I agree.

Suzanne Theberge: Great.

James Duncan: Agree.

Suzanne Theberge: So, we did have a couple of questions on usability. This measure is currently in use for surveillance purposes by the New York State Health Department. And, we had a couple of questions here, one being that the

measure is being used but there's no information provided on public reporting and whether there is any public reporting on that surveillance. And then, there was some questions here about the frequency issue in regards to ongoing surveillance, which we may have already covered to some extent. Does anybody have any questions or comments they'd like the developers to address or the committee to discuss at this time?

Male: I'm fine.

James Duncan: I'm good.

Suzanne Theberge: OK. All right. Are there any other issues, comments, questions for the developer, anything on the usability or the other criteria for this measure?

John Brookey: No, I think very well done. I think …

Male: Yes.

John Brookey: … that this is a very well done development of this measure.

Female: Yes, I think it's compelling data. Thank you.

Suzanne Theberge: All right.

(Sarah Reece): Thank you very much.

Suzanne Theberge: Thank you. Great. So, we can move on then if there are no further questions to the next Q-METRIC measure, which is 2802, same developer. This is overuse of imaging for the evaluation of children with post-traumatic headache, another process measure. And looking at the use of imaging in children ages two to 17 (Audio Gap) were evaluated in the E.D. within 24 hours after an injury.

Just pulling up my notes here. So there were some questions on the evidence for this measure. Do folks wish to – anybody wish to kick off the discussion?

John Brookey: This is John. Can I just back up and just ask a question for developer, because I didn't – maybe it's just buried in here, I didn't see it, but in terms of the

specs, are you looking specifically for a code that would say post-traumatic headache? Because a lot of times, you don't code headache, you would code minor – you know, head trauma this or that, and the headache portion would be buried in the progress notes. So, are the codes expansive enough to include all of head trauma or do you have to have a code for headache? That's my question because it makes a lot of difference as to what I think about the measure.

Michelle Macy: Yes, this is Michelle Macy, a pediatric emergency physician with the Q-METRIC team. We do recognize limitations of the ICD-9 coding, and that's why we've included ICD-9 codes, general headache, without qualification for trauma, post-traumatic headache, and also concussion in terms of the administrative claims (scoop) of cases and then require the chart review for the identification of headache with trauma, within the 24 hours preceding the emergency department visit.

John Brookey: But with head injury or head trauma in of itself, would it be included or would those be not reviewed.

Michelle Macy: Could you re-clarify that question? I'm not sure I followed it.

John Brookey: A lot of times, especially pediatricians, will just code for head trauma or minor head trauma, they won't include headache in the code. So I just wanted to know – because it – I just wanted to know how wide the (net cast). I'm trying to find out which of these kids got imaging.

Michelle Macy: So the ICD-9 code net was cast to include those cases and then the chart review was the location where we looked for symptoms of headache.

John Brookey: OK. That's what I wanted to know. Thank you.

Ricardo Quinonez: Hi, this is Ricardo Quinonez. I think this – of the measures that I've reviewed, this is one of the ones with the strongest basis in evidence. I think the PECARN Study is the definitive study on this that limiting the number of head CTs with kids with minor head trauma is the right thing to do.

If I had to stay anything is that I don't like the numerator because I don't think it goes far enough why it's limited to kids with post-traumatic headache. Shouldn't it really be all kids were evaluated for minor head injury? Whether they're in the E.R. or the pediatrician's office, anywhere, minor head injury almost never requires imaging study. And so, I'd like a little bit of clarification why specifically they would have to have a post-traumatic headache.

Male: Yes, I agree with that.

Jeffrey Susman: Yes, that was my question.

James Duncan: And this is Jim Duncan. Yes, I mean, that sort of a deviation from PECARN then. And it also at the other piece of why exclude under two years old, because I mean a lot of – I mean, there's – that's a pretty high frequency of children presenting with head trauma in that age group.

Male: Right, right, right.

James Duncan: Which is, again, part of PECARN.

Ricardo Quinonez: Agree.

Michelle Macy: So this is Michelle Macy again from Q-METRIC. The assignment that Q-METRIC received from AHRQ was specific to the overuse of imaging with headache. We felt that this clinical population was significant enough that it was worthy to address the post-traumatic headache scenario, and are in full agreement that the PECARN decision rule goes below the two-year old age mark. But we felt that children younger than two are less likely going to be able to report a headache and that's where the age breaks the PECARN decision rule.

Female: It makes sense.

John Brookey: Yes. And I think that's a limitation because for two reasons. One is I think you're excluding a huge number of patients that could be subject to imaging unnecessarily. And secondly, the way it's read and I know I'm skipping

ahead, it does require a lot of chart review to dig out that headache history. And I think it could be made to be a stronger measure that would – and require a less chart review if we were more inclusive.

Ricardo Quinonez: I completely agree with that statement, I think this measure could be (considering) can capture a lot more overuse if we don't limit it to kids who present with post-traumatic headache.

James Duncan: I agree.

Michelle Macy: So this is Michelle Macy again from Q-METRIC and we're in agreement with this being a sensitivity of the specification based on our assignment from AHRQ and are open to future enhancements to include that broader population.

Suzanne Theberge: Great.

Jeff Schiff: This is …

Suzanne Theberge: Go ahead.

Jeff Schiff: This is Jeff Schiff. I just want to say one thing in support of the inclusion of limited MRI because although there's a safety issue related to – there's also some safety in family stress and expense issues by switching even though the technique has – doesn't have ionizing radiation. It was good that that was included.

John Brookey: Yes, that was my comment was that, there really is a safety issue, a stronger safety issue with CT and radiation, MRI even though they're maybe from safety issues with the patient and so forth, it's more about cost. And so, I'm not opposed to including it, but I think there really are two different issues, right?

Jeff Schiff: I guess my concern was that the issues you put family (screw) a certain amount of expectation or you create – and we keep on raising the bar for invasive (Audio Gap). If we say, "OK, this is safe" but it's still pretty invasive to toddler, your toddler off to this just to make sure, you know, when there's

really no indications. So I don't – so I think it's good that that both are included as for an overuse measure.

Female: I think the – I think I agree that the sedation risk can potentially be an issue as well and a cost.

Jeff Schiff: Yes.

John Brookey: I don't …

Male: Great.

Suzanne Theberge: Is there anything else that the committee would like to discuss on the importance criteria, the evidence, the gap. I think …

Male: No.

Suzanne Theberge: … (inaudible) talking a bit about the scientific acceptability or some other questions about the numerator or denominator and some of the other discussions that we've had, but I did see from other – one question that was flagged a few times was regarding operationalizing abuse and neglect, or having to specify a risk of child abuse and how you exclude that. Folks, do want to discuss that?

Jeffrey Susman: Yes, I mean, I just find that – well, certainly, the willingness to document that concern has increased. I think it might be hard to pick out without rather extensive chart review. And even after that, there's probably less than 100 percent documentation. It's probably not, you know, a huge influence on this measure, which I still think as very important, but.

Jeff Schiff: This is Jeff Schiff. I guess I would think that from the E.R. perspective where I practice, child abuse and neglect is pretty well documented if it gets reported. If it's just thought about it, it's not reported. But I think it's – I think we're going beyond that stage, which cases, was there another finding that was reported.

John Brookey: Our experience actually is that even for child abuse cases that are reported and we know about 100 percent of them, we have very poor compliance of getting it into the coded section of the record. That's just our experience.

Can I ask a question about process because weren't voting on the measures today, but let's just say that we said, you know, this is really an important issue, I think we all agree it's very important issue, but we like the measure to be better. How much of a delay would that cause in getting a measure out there? I mean any thought about that?

Suzanne Theberge: Yes, it kind of depends. It depends on how much you were asking the developer to change the measure. We have a period of time after the meeting and before – after the meeting, the staff is bringing up the report on the committee's activities and then they put the measures out for comment. We can offer – one decision that you all can make is, you know, defer the decision. And the developer would then have a couple of months to do some additional testing or additional work on the measure. And then bring back some data or advice measure for you discuss on the post-comment call. (Inaudible) and (blanking) out on at the moment but I believe is scheduled for the end of February or the beginning of March.

So, if the developer can make those changes in a pretty rapid fashion and bring that back to you for the post-comment call, then it does not ultimately delay the measure. If they require more time, then it could really delay the endorsement of the measure because, you know, they would have to make their changes and however much time that would take, and then there would have to be an appropriate NQF project happening for the measure to be reviewed in. And, so, you know, that could – it could be pretty soon, it could be a while.

(Crosstalk)

Female: Yes, basically, you have to vote or focus on the measure that's before you at the meeting and …

John Brookey: Right.

Female: … if the developer can quickly provide test data, because that's generally the biggest hurdle, is the testing data.

John Brookey: Right. Right.

Female: And in my experience, that's been very difficult for developers to do in a short amount of time.

John Brookey: Yes, I just think it's unfortunate that AHRQ gave these directions, but anyway, that's my own opinion.

Female: Yes, no. And I think that's something that, you know, in an annual update, they could come in. But I wouldn't want to mislead the committee to think that there would be an opportunity to improve in a real-time cycle.

Suzanne, are you on mute?

Suzanne Theberge: Yes. No. (One problem here) to see if there re any further questions, specifications, the reliability, the validity.

James Duncan: This is Jim Duncan. I have – I really have tried to go through it a couple of times, but the denominator is difficult for me to get my head around. It seems like consist of children that were imaged with CT, is that right, rather than those presenting with head trauma that weren't imaged.

Male: You're right.

Michelle Macy: So this measure has been specified to look at overuse. And to get to overuse, we need the population who was tested. And then, within that population that was tested, identify those who received the test without an appropriate indication. So you are correct that (the children) who received imaging of their head be it through CT or MRI.

James Duncan: But doesn't that also unlink it from PECARN which is really – you know, my understanding of the PECARN study is, is that – was assessing the utility of imaging in a population with minor head trauma. And the vast majority didn't – well, don't need imaging but our image. And so, to me, that's overuse of imaging if you don't need any imaging at all but you get imaging.

Michelle Macy: So it is looking – our numerator gets us to those children who would be by PECARN exclusion criteria not indicated for imaging. And so we are getting at that population without having the broader denominator that PECARN has of all children with head trauma.

James Duncan: But, you know, talking to E.R. physicians about, you know, their use of PECARN, I don't know if they really would calculate it in the same way and use the same numerator and denominator, which I see was part of what I expected when I started looking into this measure.

Ricardo Quinonez: Yes, this is Ricardo Quinonez. I also agree that the numerator seems wrong.

James Duncan: Numerator or denominator?

Ricardo Quinonez: The numerator because if the denominator is the number of kids in which a CT or MRI was obtained in properly, then the numerator should be all children who were evaluated for minor head injury, not just who were evaluated for minor head injury and were (CT'ed) or (MRI'ed). Do you see what I'm saying? The numerator needs to be much more inclusive because, if not, you're only looking at kids who were imaged in both the numerator and the denominator.

John Brookey: Kids with headache.

Ricardo Quinonez: Right.

James Duncan: All right.

John Brookey: Post-traumatic headache, yes. So that – yes.

Michelle Macy: So, again, this is getting to that point of a difference between a used measure, which would like PECARN, all kids with minor head trauma, and then the percentage of them who undergo imaging. We are looking at overuse …

Ricardo Quinonez: Overuse …

Michelle Macy: … that denominator requires us to use the population of children who have undergone the testing in the denominator. You can't overuse without having used.

James Duncan: This is Jim Duncan. I would agree. I would say, if you just imaged everybody that comes in, that's 100 percent when only by PECARN criteria, only 20 percent which is a typical pediatric emergency room, you know, by the PECARN algorithm weren't imaging and you've got 80 percent overuse. Right, if you're scanning 100 percent of every kid that comes in and algorithm suggests only 20 percent need imaging, isn't that 80 percent overuse?

Michelle Macy: I would agree with that quick calculation.

Ricardo Quinonez: Yes. Yes, but – this is Ricardo Quinonez again. But again – and then the numerator, I agree completely with the denominator to assess overuse, it has to include all kids who were imaged. But the numerator has to be all kids who were evaluated for minor head injury, because if not, then you're only looking at those – you're not really looking at overuse if you include only kids who had a CT or MRI performed in the numerator. Does that make sense?

Jeff Schiff: I'm not sure I'm following you.

Female: Yes, I'm struggling a little bit, too. Those comments were very helpful because I really struggled with this when I was reading through the measure as well. So, thank you.

John Brookey: The other just minor thing is, is the timing going to be a problem with in terms of being able to produce the measure? I don't have it in front of me. There's a 24-hour period, right?

Ricardo Quinonez: Right.

John Brookey: Is there anything in the specs that are going to make it feasible to make sure it's put in the window?

James Duncan: Doesn't the – this is Jim Duncan again. Doesn't the specification say that they're presenting with the post-traumatic headache but I don't – I didn't see

anywhere in there where it said that the imaging is also done within that 24-hour window. And that – I guess that would be a point of clarification.

John Brookey: Yes, within 24 hours, so is it – that may be difficult in terms of – because the timing is historical and that would also require chart review, right?

James Duncan: And I also think that if the child has a headache a week after, their trauma, that would actually not be overused image then.

Michelle Macy: And that is why we went with that 24-hour time window for the emergency department visit that's paired with the imaging study.

James Duncan: Paired with the imaging, OK.

Michelle Macy: And that aligns with the PECARN time frame that the evidence to support this measure was derived from.

John Brookey: That probably would require chart review, right?

Michelle Macy: Absolutely.

Ricardo Quinonez: Yes.

John Brookey: Yes.

Ricardo Quinonez: Yes …

(Off-mike)

Male: Well, I think …

Jeffrey Susman: Wasn't exactly usability issue, right?

Ricardo Quinonez: Again, I – this is Ricardo Quinonez again. I go back to the fact that to be able to capture overuse in the numerator, you have to put all kids who were evaluated for post-traumatic headache, not just who were evaluated for post-traumatic headache and were imaged. That goes into denominator. Otherwise, you're not capturing overuse.

Jeffrey Susman: Well, it's all the ones who were obtained in the absence that documented reason to do the test.

Ricardo Quinonez: That's the denominator.

James Duncan: And this is Jim Duncan. But by changing the – you know, that side of the equation, I think you've broken the link back to the PECARN data because the PECARN doesn't, you know, calculate it that way.

Jeffrey Susman: Yes, I don't see the logic in that. I'm sorry.

James Duncan: OK.

Jeffrey Susman: Maybe I'm just being dense, that's often the case.

Ricardo Quinonez: OK. Maybe it's a better discussion for the in-person meeting.

Jeffrey Susman: You know, I mean, you're taking all the kids that have the headache and an image. And then you're looking at the ones that don't have a valid indication.

Male: Right.

Jeffrey Susman: And that, to me, it isn't specified exactly the same way. I agree, is the PECARN data. But it gets to overuse as opposed to appropriate use if you would.

Female: Yes, that makes sense.

PECARN: Does that help any? I mean as a developer, does that make sense? I mean, is that …

Michelle Macy: We believe that you're getting correct, the distinction between overuse that we're capturing in this measure and utilization of imaging for kids with head injury.

Jeffrey Susman: Yes, it's just a different way to specify similar but slightly different concept. It's overuse concept here, not use. So I believe that the denominator and the numerator are correctly specified. But, you know, it'd be good idea that we

were all were on the same page with this because it won't be easier when we have a group of however many people, 16 or 20 looking at all these things.

James Duncan: And I – this is Jim Duncan again. That – as it rolls out to 5,000 emergency rooms around the country or acute care settings, the clearer it can be, the easier it is to understand, I think, is part of usability.

Female: Yes, that's a good point.

Jeffrey Susman: Yes, yes. No, it's – I mean if this were approved, I think, you know, some real careful discussion and framing about why the measure is framed this way as an overuse measure, and why and how it's linked to the PECARN data. And while you're reading and understanding the PECARN study, may have you thinking about it this way, the reality is, in this particular measure, we're just looking for overuse. It's a little funky.

James Duncan: Yes. And I guess it comes back to, you know, how different definitions of overuse and the – you know, clearly one of the threads in imaging is overuses, you don't need a test but yet it's done. And so you take somebody that, you know, had symptoms that somebody might image when you image them. And they, you know, there's no value added. That is, you know, close to my working definition of overuse.

Jeffrey Susman: Yes, no, I hear you and I understand what you're saying.

John Brookey: I think that having been on other NQF committees, this is John, I think that, you know, the discussion might be more or less about, you know, perfect being the enemy of good. That if we – I think we all would like this to be more inclusive and to get more towards head trauma in general. But if we – if this was not approved as it may delay a couple of years getting a better measure, and so we may have a discussion about, is it better to put something out there that's going to get at least part of the population that will undoubtedly affect change and discussions in hospitals all over the country? Or should we wait for the perfect measure that may create quite a delay in getting it out there. So, I'm thinking that might be some discussion we might have.

James Duncan: Yes – and this is Jim Duncan.

Suzanne Theberge: Yes.

James Duncan: One piece – you know, again, looking at some of the data, I mean they had a pretty large – I might be jumping ahead some of the testing that they did, but boy, it comes down to a really small segment of the population that this – by the time they filtered and then did their chart reviews, it's extraordinarily small segment, right? I mean they start with a data set of, what, 60 million lives and only find a small number that actually filtered correctly and then that drops down to (Audio Gap).

Jeffrey Susman: You know, that – to me, this is all the issue around the sort of feasibility and usability. Whereas the way you were talking about the measure before and from PECARN seems like it would capture a much broader population and perhaps do more good, that this was taking a very large sample and really reducing it down to a very small number when all the exclusions were (advised).

James Duncan: And that just makes it hard to use on a daily basis. If you're only going to have the supply to a patient that you see once, you know, a year, it's not as useful as something that, you know, for the – and again, that's part of the reason to include the less than two-year old. So, you know, it's just easier just to apply all the children that come in.

Jeffrey Susman: Yes, we can't rewrite the measure for them.

James Duncan: I know, yes.

Jeffrey Susman: You know, we got to go with what we got, so.

Suzanne Theberge: That is the case but we will be having a conversation at the meeting about gap areas and areas which the committee would like to see measure developed in the future. So, you know, we can definitely highlight this as one of those areas that needs some more measurement.

Michelle Macy: Can I address the concern with the attrition in the sample?

Male: Sure.

Suzanne Theberge: Yes.

Michelle Macy: So we went from a population of 6 million lives included in the database. And then once we get into the population with post-traumatic headache, general headache, or concussion, we are down to little under 6,000 children between two and 18 years of age. And then when we looked at the number who had a CT or MRI, that gets cut almost in half to 2,967.

And then, as we get down into the sample that we used for the chart review, we were only sampling from 200 charts due to some of the feasibility issues with testing across multiple measures. And so, I think in true execution of the measure, there's going to be potentially a larger number that would be identified as potentially eligible and then giving us a bigger sample size for calculation in the future.

Male: This is just for testing.

Michelle Macy: This was just for our testing purposes.

Jeffrey Susman: But you say the majority are excluded, is that correct? Of the charts that, you know, (received that), a sample 59112 and the imaging cuts it down to 2419, but the majority of the 204 were indeed excluded.

Michelle Macy: And one of the primary reasons for exclusion of those cases was the lack of evidence of trauma occurring in the prior 24 hours. And that being an eligibility criteria for the PECARN study we felt was a reasonable inclusion criteria for this, and we are at this time without an ICD-9 or 10 code that identifies the time frame for prior trauma.

Jeffrey Susman: Gosh, I mean, I'm getting less and less enthusiastic, I mean, listening. I mean, I think the concept and the idea is great. It just seems like, "Wow, what a lot of work."

James Duncan: And didn't I read this right that the (HRD) data, it says on page 32 that it encompasses 60 million lives, which would be a fifth of the U.S. population approximately.

(Kevin Okalski): Hi, this is (Kevin Okalski) from Q-METRIC. Those lives, of course, are reflecting all enrollees, adult or otherwise.

James Duncan: Right.

Male: Yes.

James Duncan: But, so you actually were looking at a fifth of the U.S. children then if you had a fifth of the population with 60 million lives.

(Kevin Okalski): We are looking at a subset that are continuously enrolled during that period. So it's going to be less than that.

Jeffrey Susman: And even if it's a tenth, I think, you know, your point is well taken.

Gary Freed: Hi, this is Gary Freed from Q-METRIC. I think the issue is, is that this is actually how often this type of situation occurs. And the real question is whether or not this is, I think, important enough to be able to or to want to address the issue of overuse of imaging in these children (and all to make) about whether or not this number of cases that you are perceiving is really – it's important to make the quality measure for the benefit of these children and to prevent overuse.

James Duncan: I'm a huge proponent of reducing overuse. This is Jim Duncan. Again, I just wonder if the measure is structured as well as it could be to try and address this problem.

Jeffrey Susman: I mean I can well imagine the measure that is very similar but is much more useful to moving the needle around overuse of head imaging. I'm not convinced at this point that this measure is it.

John Brookey: Yes, I think we're probably on agreement that we might be able to write a better measure. But, having it tested …

Male: Yes.

John Brookey: … and having it tested and going through the process may take quite some time.

Jeffrey Susman: Oh, yes, yes.

John Brookey: I think that really is going to be the discussion. I think – I mean I just really, really would like to get us something out there.

Male: Yes.

John Brookey: I'm concerned about how this is going to go down with the committee. But maybe that's a larger discussion with the whole group.

Female: Yes.

Jeffrey Susman: Yes, I think, you know, we're at least now all in the same page in whether we agree ultimately whether this should be approved or not. I think we can usefully discuss the measure when we get together.

Suzanne Theberge: So, we just have a couple of minutes left before we have to move on to the next measure. So, I think we can discuss this further at the in-person meeting in a couple of weeks. But I wanted to see if folks had any comments or thoughts on feasibility and usability. I know we've touched on that in the discussion that has been going on. But are there other any final thoughts that you want to raise or questions for the developers before we go to the next measure?

Deborah Fattori: This is Debbie. The only thing I think that we have discussed is that the data collection process I think would be extremely labor-intensive with needing to do so much chart review, if not all, the elements were available in electronic form. So, operation would be – being able to operationalize might be a significant challenge.

John Brookey: I agree. The only thought I have to the contrary is that there (Audio Gap). But I do agree, it's a very intense chart review from this measure.

Jeffrey Susman: No, I think we've covered the waterfront now.

Female: Yes, (I agree).

Suzanne Theberge: Great. Any final thoughts or final questions in the last couple of minutes before we move on to the next measure?

Jeffrey Susman: Thank you very much for being gentle with us.

Male: It was a privilege. Thanks for your interest in the measure.

Suzanne Theberge: All right. Thank you to the Q-METRIC team. We are now going to move on to the fourth and final measure of today's call. That is measure 2820, Pediatric CT Radiation Dose from the University of California, San Francisco.

(Rebecca), are you on the line?

(Rebecca): Yes, I am.

Suzanne Theberge: Great, thank you. So, this measure looks at radiation dosage, it's an outcome measure. And, there actually are some questions about whether it's an outcome or a process measure. And I think that might be a good place to start the conversation.

Anybody want to kick that off? What we're seeing here is a measure with two components. And how do folks feel about that? There is the …

Jeffrey Susman: Well, you know, I'm not sure whether we're really trying to look at a composite measure that has two components. You know, the information, I believe, suggest that we would look at each component part in addition to the whole. And …

John Brookey: You have to get the first one, though, to get the second one, right? So you have to first measure it and then you have to look to see whether or not they exceeded and …

(Crosstalk)

Jeffrey Susman: Right, exactly, exactly.

John Brookey: So kind of in sort of a composite thing because you have to first measure it before you can see whether they're out of line. But, you're right, it could be – you could – you want to know how many people are measuring, yes, so that's important to know, right?

Jeffrey Susman: Right.

John Brookey: And you want to know how many people – how many are out of guideline, or out of benchmark.

(Rebecca): I'm not sure I see it. I didn't intend it to be two separate assessments. I thought the measure calls for looking at your doses. And that there'd be two points of comparison. One point would be comparing your averages and one would be comparing your high doses. So, I mean, yes, there are two parts. One is you have to measure those, that's basically the measure, and then second the comparative standards or benchmarks.

Jeffrey Susman: It's just the way the measure is framed, I guess.

John Brookey: Can the developer just state exactly what they mean, how this is to be reported?

(Rebecca): Yes, absolutely. So, currently, it's not typical for institutions to look at the typical doses that they use for CT scans in their patients. I think that's going to change with some new Joint Commission requirements in the hospital setting. But basically, there's more of a tradition of looking at doses for individual patients rather than looking at doses for your group of patients. So the measure calls for looking at doses, basically, for all your patients who underwent head CT who were children or all the doses who underwent abdominal CT for children. And then once you look at those doses, you'll calculate the distribution of those doses and then compare the mean on the 75th percentile.

And the reason to compare both is they both are areas that could lead to substantial quality improvement. So you could have a mean dose which gives

you different information and the benchmark doses. So looking at both together in concert, I think, provide quality improvement. If you can do one, you can do the other. So they go together. It's not like you can do one or the other. If you can quantify the distribution of your doses, then you can compare any point on that distribution.

John Brookey: So you would report it out in the two measures and not as a composite …

(Off-mike)

(Rebecca): Yes.

John Brookey: Right. My one question is about – I think that this is really, really important. I see that this is really right for opportunity and what the right timing as well of the Joint Commission. How reliable from institution to institution are these measurements? And is that really that relevant if you're willing to just try to get overall improvement?

(Rebecca): So, on a practical operational level, there are two ways that an institution defined as a single facility, a group of facilities, a health plan, could assemble this data. They could assemble them using an old fashioned approach, where they basically sit down at a terminal and write down the numbers. And we wrote a paper checking that old fashioned approach. But you can imagine institutions where there would be resources involved in looking at the numbers. And if they do very few scans, it might be easiest to do it manually. So we did at that approach, but my expectation would be that most institutions would do an automatic approach. And basically, either use – there are a whole bunch of tools you can do this with now, but dose software or pulling the data from the (PACS) or the radiology information system, and that automatically, basically collates the data.

And we've done it both ways. And we've checked it both against, you know, sort of reference standard of sort of pulled manual review. And the data are highly reliable, you know. As you'd expect, there are a few errors in writing down the numbers when it's done manually. But the impact of the summary statistics is negligible. We did it in the context of several projects to make sure that the tools we were using to automatically extract the data were

actually doing what we thought (Audio Gap), and the results were highly reliable. So, you got what you thought you were going to be getting.

Jeffrey Susman: Another area that, I guess, I have questions or concerns, comments about is the idea that measurement will drive improvement. And, you know, I'm always skeptical that just measuring the (phenomena) will actually drive improvement.

(Rebecca): So, I think that is a very important point. And just to give a little background, the measure had been previously endorsed in a very similar format and that unendorsed because of the question that you're raising, which is really is just looking at your data and comparing to benchmarks enough to drive practice. So, I've included in this new submission, which is quite similar to (inaudible), with some very small modifications and improvement. We did just that as we did several projects. We've now published several and some are several on press, where we basically assembled these data, gave them to the facilities to look at. So, just looking at their data and that drove pretty substantial improvement depending on which project between 10 percent to 30 percent reduction in doses, where the only really comparison was, how you were doing compared to other benchmark.

Jeffrey Susman: And these were around audit and feedback educational programs?

(Rebecca): Yes, that wasn't really so much more than just the audits that these are your doses. This is how you compare. And we saw substantial decrease. We actually have a large grant now, which is a randomized trial at a hundred and – somewhat 125 hospitals in the U.S. and abroad, where we're comparing simple audit just looking at the doses from more multi-component intervention that includes a lot more education, a lot more access to expertise, understanding institute of health.

I agree with you that there's probably ways to make the feedback more robust to get a greater improvement and we're starting that stud. And you have the simple audit just looking at your 75th percentile and your 15th percentile had by itself a substantial and significant reduction in doses.

John Brookey: Can I ask a naive question? Isn't it typically the tech that's determining the dose based on protocols for that hospital or institution?

(Rebecca): So, I don't think it's helping me to confirm your naivety. But it ends up incredibly multi-factorial how those doses get set. And they involve the technologist, the physicist, the radiology, residents, administrators, and (inaudible) (at much the same). So, at some institutions, the technologist is entirely responsible. In some institutions, technologists have almost no role and whatsoever and it's everywhere in between. So, I have a lot of thoughts on other ways to optimize the doses that are used. But in fact, I would guess the technologist is a very important player. But by no means, you know, the only player that influences this. Actually, it's a really complicated process, and one that's not standardized. So every institution does a different one.

John Brookey: I've already engaged our physicist in the conversation, and then they're ready to go, so.

(Rebecca): Excellent.

Suzanne Theberge: So, any further – I think we kind of have jumped here on a little bit, but any further questions about the importance of the gap or the evidence of the measure?

John Brookey: Can I clarify? Does this go to a previous NQF committee? Or was it previously not approved?

Suzanne Theberge: That was in the patient safety project.

John Brookey: Oh, OK. And how long ago was that?

Suzanne Theberge: (Rebecca), correct me if I'm wrong, but I believe that was 2014, late 2014?

(Rebecca): I think like that. It was approved its first submission. It was active for some period of time. And then the next submission really (asked), and I could not provide evidence that actually looking at your data could have any impact on dose. And I'm turning it around so quickly because we've just had a lot of success but that's precisely that. Yes, looking at your dose and comparing it

does result in lower doses, not as low as I think they need to go, I don't think it's the only way one has to do it, I think, you know, educating a technologist is an enormous need as well as other things but I think by itself. So, I'm turning it around quickly. And the prior submission was people of all ages, and this one is focused on children because I think the impact of the radiation dose and overdosing is potentially harmful in children so that's why it's in this group.

Suzanne Theberge: Yes, I think the measure was originally endorsed in 2010 or 2011, if I'm remembering it correctly. And then came back to patient safety about a year and a half ago.

John Brookey: Are there any potential downsides to this in any unintended consequences?

(Rebecca): So, I think there is a theoretical downside. And I say it is theoretical because I sort of feel like we should be so lucky for that to happen. But, at some point, if there's a lot of external pressures to lower the doses, you could lower them too low to the point where your exams are no longer diagnostics.

Now, of course, it's a responsibility of the radiologist (who gets it) and say, "I can't read these images," because in general, the higher the dose, the prettier the images. And it's the job of the clinical radiologist (Audio Gap) doses down to a safe level but not bring them down too low. And so if they're too low, the radiologist should be complaining. And there are a lot of systems in place that actually help that to happen. But that is a theoretical unintended consequence. But we have found in our (Audio Gap), often orders of magnitude higher than they need to be. So I think there's a lot and lot of room to bring them down to a good job before when the (zone worth) is too low.

I didn't include in the measure looking at low doses, right, making sure your doses don't go between too low because currently there's such variation, there's no way really to set that threshold in a way that wouldn't have an unintended consequences penalizing people who are actually doing their best job because their doses were so low.

John Brookey: I know we have two radiologists on the line, and one of which said he’s going to recuse himself. But I’d like to hear their comments.

Suzanne Theberge: Dr. White, are you on the line? Do you have any comments to share?

I know he has to step away for another meeting. He may not be back yet. But he will be on the line for the in-person meeting.

James Duncan: This is Jim Duncan, and I'm sorry …

(Crosstalk)

Suzanne Theberge: So, if you folks have further questions on, you know – and I think we've actually addressed a lot of the criteria so far. We haven't talked about feasibility. Any questions about feasibility? Any comments about feasibility?

John Brookey: Well, every imaging center has the ability to measure dose, right?

(Rebecca): The measures that I'm calling for looking at are reported essentially by 100 percent of the CT scans in the country. The number might be 99.5 percent. So, yes, measures that I'm specifying are on and collected for every CT scan.

Jeffrey Susman: So, is this an intermediate outcome measure or clinical measure? Or is it a process measure. I'm not …

(Crosstalk)

Jeffrey Susman: What's that?

Female: That's what I've struggled with, too.

John Brookey: If your outcome is just to have less radiation, you could call it …

Female: An outcome measure.

John Brookey: Or of an outcome measure, if your outcome is to have less cancer, then that's – this is a proxy for that. If your process is just to provide the appropriate dose, then that's a process measure. So it's kind of almost semantic in a way.

Female: Yes.

Female: True, true.

Jeffrey Susman: I mean, one way to look at it is does the patient really care how much radiation they get. I mean, yes probably, but.

John Brookey: Not until they get cancer.

Jeff Susman: Right, exactly.

Female: Yes, I guess for me, as I was trying to (grave) the evidence that what type of measure it was, was factored into that.

(Rebecca): I mean, I – you know, it's not really for me to say, but I think if you – well, we have actually surveyed patients as part of several projects that I probably could have cited in here. Actually, one was just published. They do care about radiation dose.

Female: Yes, yes.

(Rebecca): Now, I think they care about it because of its …

(Crosstalk)

(Rebecca): … they care about radiation as an (end) in and of itself.

(Crosstalk)

Female: I think parents of our pediatric patients struggle with that. Many parents do.

John Brookey: Yes. It's probably in here, I just haven't – I don't have it open. But if the – is there some academy of pediatric radiologist, I should know this – I mean is there some sort of guidelines that are out there that basically specify, you know, the range of doses for various conditions for pediatric patients? Do we have something to benchmark off of?

Jeff Susman: There were things that were cited there. I don't know if that's …

John Brookey: Yes, I just forgot whether it was there or not.

(Rebecca): So, there are different groups that are publishing different numbers to (that offer) that a good numbers are not good numbers. They don't vary a whole heck of a lot. They're relatively similar. But I think there's need for more benchmarking. So, you know, that's part of the …

John Brookey: Right.
]

(Rebecca): … measures (as compared to) benchmarks. And we have published some benchmarks (with) several other larger papers that will be coming out soon. There's no single that, you know, what's a normal lab value, what's a normal radiation dose. They're relatively wide ranges.

But, if you look at children's images, image in children's hospitals versus children imaged in adult hospital, the doses that they receive are very, very discordant.

John Brookey: Yes.

(Rebecca): So the doses that you get in a children's hospital just because it's much greater sensitivity to this are no surprise much lower.

Jeff Susman: Right.

John Brookey: And they also have better expertise to – so this reminds me just a little bit about our efforts to not use contrast routinely in abdominal scans. So, it's sort of an efficiency measure in a way and not all the radiologists are onboard with it because they know they're not going to get a good of an image without the contrast. But they're willing to …

Jeff Susman: Right.

John Brookey: Do you see what I'm saying is that I – it's – that's why it's really important for me to know how much – how onboard the radiologists are going to be and is a pediatric radiologists at an academic institution going to really be able to image with lower radiation cost of their expertise compared to a community hospital where we often don't have pediatric radiologists. So that's kind of my getting to the unintended consequences question. But it sounds to me that the

measure in a way is just going to inform all of this. If everybody measures that in reports that their means and comparison to benchmark is kind of going to probably move organically in one direction, which is going to end up in lower doses. I think that's kind of where this is going. But we may not have every radiologist who might be completely onboard.

Jeff Susman: Yes, I think the protocols that are used around, for example, single versus multi-phase study might be moved as a result of this. I mean one could envision if we believe providing the information will change behaviors. Will there be at a plan level, for example? Oh jeez, our facility looks really bad compared to, you know, with other facility or this plan looks worse than another.

John Brookey: Right.

(Rebecca): That would be my hope. I think what you focused on is the use of multiple phase study, the single most important thing a facility could do to lower their doses relatively quickly. And it's the use of multiple phase study doubles the dose or triples the dose, or quite triples the dose but the choice of those study tends not to be evidenced based, it's more of (inaudible) your pictures, it's interesting we're doing that. But the only way to sort of motivate, I think, doing less of that is to look at the doses and realize, "Oh, it also (leads) some doses that are …

(Off-mike)

John Brookey: Right. Right.

Jeffrey Susman: OK.

John Brookey: I think this is going to be a really interesting discussion for the whole group.

Suzanne Theberge: Yes. And we're …

Jeffrey Susman: (Are you guys ready)?

Suzanne Theberge: Oh, go ahead.

Jeffrey Susman: Oh, I was just going to say, I think I've heard enough on this to be dangerous.

John Brookey: Yes. I'm going to just do some talking to our pediatric radiologists before the meeting but I mean, face value, it sounds really, really excellent.

Suzanne Theberge: All right, are there any further questions, last minute comments or questions for the developer before we move on to close on the call? We just got a few minutes left.

(Rebecca): It was just something that I'm not sure if maybe I just filled out the wrong boxes that I thought I could clarify about exclusions or not exclusion.

Female: Yes, that would be helpful because I was a little confused by that.

Suzanne Theberge: Yes, we just see some comments about that. And if you could, -- (Rebecca), if you could do that very briefly now and then we'll follow up with you to make sure that that gets shared with the full committee in writing.

(Rebecca): So, the measure calls for looking at every exam that's done in the most common anatomic areas in children, so looking at consecutive head scans that were done for diagnosis. It doesn't call for looking at CT scans that were done as part of biopsies or procedures, it doesn't call for looking at scans of other anatomic, it doesn't call for looking at scans that were done as part of radiation oncology.

So, it's inclusive of what it focuses on. And then, in the exclusions, I included areas that happened not to be part of those anatomic areas or those indications. And I'm not sure if, you know, that's quite the right way to word it. But we're just not looking at spine CTs, we're not looking at biopsy CTs. They will introduce much more variation in dose. Maybe one day, they might provide their own good measure by themselves but I didn't want to make the data on the simple anatomic areas less precise.

Suzanne Theberge: OK. We'll – NQF will follow up with you after the call, and we'll get that clarified in new submission for the full committee to have prior to the meeting.

So, at this time, we're almost to the end of our call, and I'd like to pause here and open the line for public comment. So – and also if you haven't dialed in, if you're just listening to the streaming audio, then, you know, you can also just comment via the chat box. But operator, can you (open the lines)?

Operator: Thank you. At this time, if you have a comment, please press star then the number one on your telephone keypad. We'll pause for just a moment.

And there are no public comments at this time.

Suzanne Theberge: OK. And I don't see anything in the chat box. So we can move on to the next slide, which is the next steps. So, the next steps are, we have now finished workgroup call with this call. We'll be posting the recordings and transcripts of the other three calls and this call up on the web so that you can take to look at the other measure discussion.

This time, if you have not already done so, we'd like to ask you to review the remaining 11 measures in this project. So, Dr. (Gin) will look at this more in the worksheet. We are taking all of your comments that you all submitted via surveys and putting those into the measure worksheet that we posted on SharePoint. And we're going to have this updated forms on SharePoint tomorrow. We'll send around an e-mail when we've got those up so that you can see what your colleagues (have said).

And either tomorrow or Friday, we're going to be finalizing the lead discussants for each measure and we'll be sending a lot more information about that. But basically, we're going to ask each of you to ask as a lead discussant on two or three measures. We will have two or three lead discussants from each workgroup on each measure in the workgroup and you – lead discussants will be responsible for kind of kicking off the discussion, raising the point that came up on this workgroup call. It's used in the submission form by NQF staff, et cetera. So, (we're going to be charged) with leading the discussion on your measures during the meeting.

And travel arrangements because we are meeting in two weeks, we will be seeing you all in Washington, D.C. And so, you should have your travel arrangements squared away. You should have received your hotel

confirmation information for our meeting team. And if you have any questions or concerns about that, please definitely e-mail me.

Next slide. So, as I said, we'll be meeting in a couple of weeks. And then, if we don't get through everything at the meeting, if, you know, we get to the end of the meeting and we just haven't finished discussing all the measures, then we will have a follow-up call on December 10th and we'll know by December 2nd whether or not we need that call. And I will pause here and see if there are any questions.

All right, well, hearing none, next slide. We will have our contact information posted. So, if you have any questions as you go through the measures, give us a call or an e-mail. Any questions about anything really, really the process. We are looking forward to seeing you all in a couple of weeks. Thank you so much for your time so far and your time today. It's been a really great call, lots of the feedback for developers and for NQF staff. So thank you very much. And I hope everyone has a wonderful Thanksgiving before we see you.

Female: Thank you.

Female: Thank you.

Jeffrey Susman: Thanks so much. Take care.

Suzanne Theberge: All right, thanks, everybody.

Jeffrey Susman: Bye.

Male: Bye.

Operator: This concludes our call and you may now disconnect.

END