NATIONAL QUALITY FORUM

Moderator: Performance Measures Pediatric
November 12, 2015
2:00 p.m. ET

Operator: This is Conference #: 16840631.

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

Suzanne Theberge: Good, everyone, and welcome to the Pediatrics Performance Measures Steering Committee Workgroup Call Number 2.

Thank you very much for joining us today. This is Suzanne Theberge. I'm the senior project manager on the team. And like to go ahead and get the call started.

Before we begin, I would just run through our usual housekeeping announcements. Committee members, you will need to be dialed into the phone to speak and to have an open mind, we can't accept talking through your computer lines. So please dial in if you are, if you would like to participate verbally.

We also do request that you turn your computer onto mute, turn your speakers off if you are calling, and streaming the webinar because we will get feedback on the line from the webinar audio.

So if you have any questions or technical issues, please feel free to send us a chat. And for technical issues or do have a question about the criteria or anything, please feel free to just jump right in. So next slide, please. And actually, the next slide, we will jump right into the roll call.

I would like to start by doing a quick roll call for the committee and as I say your names, please let us know that you're here. And just gives one or two (minute) introduction of who you are. And going forward, we do ask that as you speak, you just say your name first so we know who is talking. And after the committee has introduced themselves, we'll have the team introduced themselves.

Lauren Agoratus, are you here?

Suzanne Theberge: Jim Bost?

James Bost: Here.

Suzanne Theberge: Great. Can you introduce yourself briefly?

James Bost: Oh, sorry. Yes. I'm Jim Bost, Director of the Outcomes and Quality Measurement at Children's Healthcare of Atlanta.

Suzanne Theberge: Thank you. Karen Dorsey?

Karen Dorsey: I'm here. I'm a Pediatrician and Director of the Division of – Measure Reevaluation and Role Making at Yale (Core) where I'm a measure developer.

Suzanne Theberge: Thank you. Kerri Fei?

Kerri Fei: Hi. This is Kerri Fei. I'm a Clinical Value Portfolio Manager at BlueCross BlueShield Association in Chicago, where I work in performance measurement.

Suzanne Theberge: Thank you. Amy Houtrow?

Amy Houtrow: Hi. This is Amy Houtrow. I'm a Pediatric Rehab Medicine Physician and Health Services Researcher of Children with Disabilities, at the Children's Hospital of Pittsburgh.

Suzanne Theberge: Thank you. Marlene Miller?

Marlene Miller: Oh, I'm here. This is Marlene. I am Vice Chair of Quality and Safety and Chief Quality Officer for John Hopkins Medicine for Pediatrics in Baltimore, Maryland.

Suzanne Theberge: Thank you. And, Lauren, did you join us?

All right. Well, hopefully she'll be able to dial in later.

So we, also, would like to have the NQF team introduce themselves introduce themselves. As I said, I'm Suzanne Theberge. I'm the senior project manager on the team. Nadine?

Nadine Allen: Hi, I'm Nadine Allen, Project Manager for this project. I'm also project manager for the MAF-Medicaid Child Project, also the Prenatal Project and the Surgery Project at NQF.

Suzanne Theberge: Severa?

Severa Chavez: Good afternoon, everyone. This is Severa Chavez and I'm the Project Analyst. And in addition to this project, I'm also in other MAF and Medicaid task forces, thank you.

Suzanne Theberge: And Robyn?

Robyn Nishimi: Robyn Nishimi, I'm a Senior Consultant to NQF, arrange of this huge part of this, I was the Founding Chief Operating Officer of this since 2007. I have moved over to a consultant role of doing more content. I am more content. I am, primarily, have been with NQF on the disparities and population health projects, but pulled into this.

Suzanne Theberge: Thank you. All right, and we also do have the measure developers on the line. So committee members as we're discussing the measures that we have on the agenda for today, if you have questions, please feel free to speak out and ask them your questions.

And I know we've also got a couple of other committee members on the line who are not on this workgroup, that have dialed into listen to the conversation. And I just like to welcome them as well. And now, next slide.

So just to go through the agenda and process, briefly, before we get start. We're going to be looking at two measures today. We're going to walk through the criteria for each measure, each of the criteria which, as you may remember, are important scientific acceptability, feasibility and usability.

We're not going to be voting on anything today. Today is just more of a conversation about the measure. And measures and issues that the workgroup raise during the preliminary surveys that you all completed.

As we go through these measures, we're going to focus on places where the workgroup members, where NQF staff raised the question about the submission. But if, you know, we had a concern with the particular area, then we're going to spend that much time on assessing it.

Following our discussion, we will move into NQF member and public comment period. At the end of the call, we'll just have a few minutes, in case anything on the line and wishes to make a comment and then, we'll discuss next steps.

And before we begin, I did want to remind folks that NQF does in endorse measures primarily for accountability purposes, but we do also endorse measures for quality improvements.

So it's not just Q.I., it's not just the accountability, so we do need to think about measures dealing accountability lens as well as the Q.I. lens, and next slide.

So I'm going to turn it over to Robyn to lead the measure discussion but I just wanted to note, we had a last minute change of plans, so we're going to discuss the ADAPT (and FECC) prior to discussing an effect due to a slight change. So we're just going to have to move things around a bit.

So with that said, I will turn it over to Robyn to begin the discussion on measure 2789, the Adolescent Assessment of Preparation for Transition to Adult Focus Healthcare. Robyn?

Robyn Nishimi: Great. Thanks, Suzanne. So what I'd like to do is mindful of the time. And that even though it looks like we only have two measures, both measures have several subparts, is when we – if it looks like we're bogging down in a particular area, then I'll interrupt the workgroups members and trying sum up where we are at that point so that we can move on to the next area.

I don't want to cut our discussion too short but I do want to make sure that we get a chance to discuss and air the major issues of both the ADAPT and the FECC.

As Suzanne indicated, we're not going to vote on this and I don't expect that workgroup needs to come, to consensus that really is more key process of giving you all a chance to discuss your specific support or concerns about a particular measure. If it's readily clarified on the call we would ask our developers and we'll also know to the developers. We asked the developers to kind of hang on the sideline for the most part. If this is something that we think the developers can readily answer. Then, we'll go to the developers for clarification, otherwise the expectation is that, the developers would be prepared to address your concerns at the in-person meeting.

So with that, as background, are there any questions on how we're going to proceed?

James Bost: We were totally would get an update spreadsheet with the correct columns, did that get sent or did I'm missed it? This is Jim.

Robyn Nishimi: No. We got – I think we got one more response from a committee member, but we were not able to get the response since we shuffle. There's a problem with the import from your survey into the NQF system and we couldn't that result.

So for the most part, what the columns that you see are correct, that there are few. I think it's generally – I went through them this morning and pasted them in different places. So I think it's generally clear where things are and I'll be summing up things for each sections. So I think …

Lauren Agoratus: I'm sorry to interrupt. This is Lauren Agoratus. I just joined the call, I am actually in a hotel room of Route 1, and I have no internet access. So I apologize for being late.

Robyn Nishimi: Thanks – you could join.

Lauren Agoratus: Thank you.

Robyn Nishimi: So the first major component for the workgroup discussion is on the evidence. This is a patient reported outcome, so the committees vote will be a pass or no pass, based on the type of measure it is.

The evidence that needed to be presented was a rationale that the target population values the measure, find it useful and that the measure is supported perhaps by other literature. But really, as the patient reported outcome, the level of evidence and the type of evidence necessary is different from you'll often see for process measures.

There was mixed support for the age range, some support in 16 to 17, some support for 14 to 15 years, and then there was support up to 21 years. Developer is asked to address the – whether they assess the value and usefulness of the survey through their focus groups.

There was concern that there was no evidence on the domains or evidence that the counseling provided actually improves readiness. In some respects, those (inaudible) validity of the specifications, but we can discuss them here.

That's what I took away from the committee input and with that, I'll open it up to the workgroup members to elaborate on their comments or disagree or agree with their colleagues.

So let's ask – let me ask the developer this. These, the focus groups, assess the value of usefulness of the survey.

James Bost: And specifically with clinician, again, this is Jim, sorry.

(Greg Stawicki): Hi. This is (Greg Stawicki) here. I'm one of the co-leads for the ADAPT measure. And to address that particular question, during the focus groups

themselves, we did not specific ask the participants in the focus group about whether they felt such a measure would be useful. The focus groups were framed around the idea that these focus groups were being done to help develop such a survey and a measure.

But I see that some of the members asked about whether clinicians found this to be useful. So we also conducted a series of stakeholder interviews and clinician interviews prior to developing the focus group interview guidance survey questions both within our institution and nationally among experts of around adolescent health and transition.

And there – although, this is not sort of a formal analysis, there was quite a bit of support around having these types of domains and types of questions involved in a patient reported and adolescent reported measure.

James Bost: So the choice of the domains that you did in terms of evidence was derived from these interviews and discussions.

(Greg Stawicki): Derived from interviews with experts, derived from the focus groups themselves as well as the review of the literature that have insist, as well as the review of the consensus statements from organizations like the American Academy of Pediatrics.

James Bost: Thank you. And the age decision, again, it looks like we have some different thoughts about the appropriate age range, could you comment a little bit about the evidence for the age range.

(Greg Stawicki): Sure. So we – this was a big discussion among our group as well. We note that the consensus statements nationally recommend transition planning and preparation to start at age 14 or even some recommendation down to age 12. We felt that that would be too young of a place to start, for youth reported measure and felt that if we had started age 14, it would at a point where many use probably wouldn't be able to answer around this type of care. So we felt that by 16 would an appropriate place to start.

We also felt that although, this – that it's your – I think that the point that this could be extended to an older adolescent and young adult group is very fair.

There is many institution around the country and certain clinics and clinical programs where they consider age 18 as the age that which they're transferring patient to an adult center. That's not universal and there's certainly debate in the field about the appropriate age for such things, to whether age would be a criteria at all.

But we felt that looking at 16-year-old and 17-year-old, this would be adolescents who should have had some preparatory work in guidance and counseling around transition already done, and they would have been uniformly in clinics that we're still being treated by pediatric providers, whether they would be subspecialist or primary care providers.

James Bost: Yes. I think the age choice often depends on the condition from my experience here.

Amy Houtrow: Yes, this is Amy. I was thinking basically the same thing. And certain programs for example, complicated congenital heart disease, those individuals might never really transition kids with developmental disabilities or other cognitive impairments might not be capable participating successfully in this kind of process regardless of age. And if there are a lot of disease specific differences.

Robyn Nishimi: OK. Any other question that the committee want or elements the committee wants to discuss besides the age and the issue of usefulness?

(Karen): This is (Karen). I was going to ask also – so I mean, in your earlier statement, you indicated the choice of age had more to do with the ability to complete the survey. And I'm just curious if you all, that was sort of purposeful consideration and the design of the questions. I mean, you certainly – you can develop a survey that younger children could complete but, you know, (inaudible) from inception, you know, worried at this – structure the questions or the content of the question as such that it just – you just couldn't adapt it for younger children.

(Greg Stawicki): That I think an appropriate comment and appropriate point. So when we did our focus groups, your youngest participants were age 16. And then, when we did cognitive interview testing of the draft measure itself, it was specifically

with 16 and 17-year-old to make sure that it was at the appropriate understanding and reading level for that target population.

Ricardo Quinonez: Hi. This is Ricardo Quinonez. Although I was not in the workgroup that was supposed to look at this indicator, I just have some comments. So I think, you know, it has a lot of face validity, you know, the – I think everybody would agree that it's important for, you know, children to transition safely and effectively to adult care. I worry that there's very little evidence on whether the intervention that's presented here would actually have any impact on outcomes.

So in that case, I see the evidence that supports it, as not very high quality. My comments on – really just looking through the evidence percentage and the early initial draft.

Marlene Miller: This is actually Marlene. I think I completely agree with that. That was the comments I put – I think nobody will argue that the concept is important, but there is almost no evidence that clinician counseling is the correct intervention to actually achieve it. And I do think we really have to consider that because we don't want just assume that this is a correct intervention that will actually achieved condition readiness.

Lauren Agoratus: This is Lauren – right. This is Lauren. I just wanted to mention that our organization supported 14 to 15, so the American Academy of Pediatrics with the age. We did this also for transition, so there were like for special education. We have that put into our state, especially at the New Jersey Administrative Code. I'm just going to throw that out there too.

(Karen): This is (Karen) again. I think, you know, another way to think about it is the validity question. I completely agree that it has a lot of face validity but the other way to think about it whether or not we consider this kind of conversation initiation of this kind of counseling as a threshold event in planning. And so, I completely agree that there's not evidence, the discussion alone is sufficient to sort of ensure smooth transition to adult care.

But, you know, might we consider this a threshold event where if you don't at least have an initiation of this conversation and there's really no hope of

having a fewer transition. So just another thought about how we might think about validity.

Robyn Nishimi: I guess – but I think that's a true point, I think the one thing that always fix in my mind thought as we continue to raise that question are these being used for quality improvement or for accountability. And the answer is usually something that we can't really prescribe one way or the other, and they may end up being used for accountability. And then that's where that gets problematic. Because then, there's a whole other machinery that gets involved in resource intensity on an intervention that perhaps is not the right intervention although we all agree that the concept is important and …

Female: Yes.

Robyn Nishimi: … to happen.

James Bost: Yes, this is Jim. I'm not sure, I quite get the concept of intervention when assessing a measure, you know. I look at these items as, are they important to understand in transition or is there room for improvement in the items. And as we get, you know, are there gaps that these items show in transition to care so I thought they did address those issues. So I'm a little not sure where we're going with intervention when we're talking about developing a measure.

Marlene Miller: Well, this is Marlene again. Well, I guess, you know, in my view what the survey suggest is that, the transition of readiness is achieved by clinician counseling on this very concrete bullet in these questions, correct? And what we don't know, I mean, what we want, what these measure is trying to – that we prepare for transition to adult.

So what we want to be measuring is the intervention that we know will best prepare – for that. And we just don't know and there's – the reason for a lot of skepticism that, you know, a one hour counseling session from a clinician once a year is going to actually achieve this and shouldn't be – instead be focusing on some other interventions or research on what would be an intervention that would better prepare. And then measure the use of that as opposed to assuming which we often do that just because the clinicians says

that that is the answer of fixing whatever that problem is in terms of the patient education or need.

Amy Houtrow: This is Amy. I agree with that. I mean the connectivity between the intervention that the provider does and the outcome of actually being ready for transition. It's like we're questioning, right?

Marlene Miller: Correct, correct, when you want to be measuring something, it actually prepares people for readiness. You don't want to just measure that we do counseling even though it does not help people be ready for transition.

Robyn Nishimi: OK, anything else on this area?

OK, let's go on to gap. There was a similar issue in performance gap. Most comments appeared to indicate that I think transitions and these measures demonstrated the gap in care. But there was one comment that, again, indicated that the submission didn't provide data comparing how adolescent believed to be well-prepared score compared to adolescents who are not well-prepared.

So again, I'm sure that the survey actually detects transition readiness and ergo it was not possible to indicate whether there was a gap in care. So if the workgroup members wanted to opine on that and then see if they had any questions for the developer.

Then, let me ask the developer about the question of whether you have data comparing how adolescents who are believed to be well-prepared presumably because they've scored well, compared to those who are not well-prepared.

(Greg Stawicki): So because our testing thus far has really only been cross sectional in nature, we only have information about the current status of individuals who completed the survey so we can't really comment. Then, I think it gets to the point that's been raised around the – right now, lack of evidence that suggest that this type of measure or for that matter any measure of transition readiness is going to be associated to outcomes further later on once transfer or transition is occurring. And I think that it's something that we weren't unable to test at this time.

Ricardo Quinonez: This is Ricardo Quinonez. Do you have any idea of what outcomes you could possible measure in the future to sort of assess the validity of the …

(Greg Stawicki): I think that's a great question and the outcomes that have been looked at in sort of smaller studies or in disease specific studies certainly a place to start. Those include things related to access to care, gaps and care, loss to follow-up, things related to medication prescriptions for those that are actually taking medications, you know, refill rates, you know, healthcare utilizations, urgent care visits, acute care visits, E.R. visits.

There is literature in asthma that shows that young adults have (higher EVUs) with asthma compared to adolescents and there is a longitudinal data suggested there is a rise among similar type of population.

So I think healthcare utilization outcomes as well as sort of access outcomes would probably be the sort of first pass in terms of looking at sort of that connectedness with the healthcare system that often is at risk on any transition period.

Ricardo Quinonez: Yes, this is Ricardo Quinonez again. Those sounds good and sounds like you guys have thought about it. Can I suggest another one? Have we thought of quality of life? There are very good validated quality of life service which you could do, you know, when there are adolescent and then when they are – they have supposedly transitioned to adult care. You would hope that their quality of life will be actually the same if not improved with the transition process that's going well.

(Greg Stawicki): Agreed.

Robyn Nishimi: OK. Does the – do any other workgroup members have additional thoughts on whether gap in care exist or any other question?

OK, let's move on to reliability then. The reliability criterion has two components, the reliability of the specifications and then the reliability of the – the reliability testing, sorry.

So, the first component is the reliability of the specification. I would say, overall, it was felt that the algorithm that was provided was clear that with the exception of one very specific case, the questions appeared to be clear at an appropriate comprehension level. There were some questions about the sampling but overall I would say that the reliability feedback that you gave were generally consistent. There were concerns about administration as indicative with the legal guardian, even to be present in caregiver assistance for those with developmental disabilities. Again, these are implementation issue.

There was a question about the term, "being more in charge of your health as being clearly defined to develop and might need to be queried about that". And with that, I will turn it over to the workgroup members to discuss any questions, or concerns, or comments that you specifically want to make to your fellow workgroup members.

Awfully quiet after that spirited evidence discussion, are there any concerns about the – whether the specifications can be reliably implemented. If not, we'll move on.

Marlene Miller: This is Marlene. I guess I'm looking at my (comms), what I don't know is whether this (PMCA) tool which is part of the denominator has been tested in multiple Medicaid databases just knowing that they are all structured very differently. Do we know?

Robyn Nishimi: Developer can you address that?

(Greg Stawicki): So this was a tool that was developed by another workgroup that's part of Centers of Excellence for pediatric quality measure development to identify youth with chronic – complex chronic and non-complex chronic condition. I think it's been tested in different Medicaid populations. But certainly at the in-person meeting, we can get back to you about sort of the degree of that testing. But it was one of several algorithms that we look at to try to evaluate way to select in a claims database in appropriate population to survey.

Robyn Nishimi: You know, yes that's what I thought too. I thought it's been tested too. But that's, obviously, and we know, I mean I was in the Medicare database and

Medicaid database, they're all completely differently structured, and if it is not feasible and then a more – much more specific to Medicaid in a very much limit the feasibility of this measure.

Rita Mangione-Smith: Robyn, this is Rita Mangione-Smith. I don't know if I'm allowed to say anything but I can certainly address this issue.

Robyn Nishimi: Why don't you go ahead because it was raised also under the feasibility criterion. There was concern about denominator.

Rita Mangione-Smith: Sure. So the pediatric medical complexity algorithm has been tested in the 50 state (Max) Medicaid database and function very reasonably in that database. It's been tested in our center in Minnesota Medicaid and Washington Medicaid. It's being used by over 20 different healthcare organizations using different kinds of administrative data successfully including Cincinnati children, Geisinger. Well, I could name several more if I had my list in front of me.

So it's been widely used and implemented in different administrative databases and validated in two or three different databases. So I think feasibility does not appear to be a problem with this.

Robyn Nishimi: Any other questions about the reliability of the specification?

James Bost: Was any test-retest reliability assessments done?

(Greg Stawicki): That was not done as part of our field testing.

Robyn Nishimi: And this is moving into our reliability testing.

There were several concerns expressed that the results weren't clear about the critical data element level testing. There was concern about the test sample due to low response from the health plan.

As Jim mentioned, some test-retest – small sample might have been called for there was a repeated testing in the same population to assess the reliability and testing was done only in three geographically dispersed areas but no

information was provided as to have similar the test populations were or weren't.

Are there any questions about the reliability testing per se? Anything you want clarified from the developer? OK.

Let's move on to the validity of the specification. We covered some of this when we discuss the general evidence. The concerns that were raise where there no real specifics on the need to split this into the three domains and why those domains we're chosen. I think we've discussed that. And then, there was no indication that the intervention counseling was representative of a good or bad transition readiness, I guess, if you will.

And then there was concerned that the survey in domains didn't address the concept of care coordination amongst multiple providers. So the measure doesn't capture that. Any other issues that the workgroup wants to make raise or rather point developer would like to respond to once the work was done?

Committee members? No? And, developers, if you want to respond to any of those points.

(Greg Stawicki): The only thing I might comment on this. We acknowledge that this wasn't really measure around care coordination. And we understand that a lot of adolescents with chronic health condition do require care coordination, but we felt that that was beyond the scope of this current measure.

Robyn Nishimi: OK, moving then onto validity testing. The developer indicated that validity testing that the performance measures score was performed. The questions that were raised by the workgroup members were primarily around the transfer planning measure to the small sample size. And so whether in fact that was a valid measure and indicator of quality. Let's see.

No data or testing were done to address whether the target populations values the concept captured by the survey as meaningful. And then there was a comment that validity system appear to be sufficient.

So there was mostly concern about the transcript planning measure and then whether they are meaningful differences and whether the concepts captured by the survey were meaningful to those being surveyed.

Is there anything that committee members would like to comment on or any questions you have for the developer?

James Bost: Yes, hi. This is Jim. I just – I might have missed it but I didn't see any specific validity testing to see if these we're similarly to similar domain type items that are out there, in other words some criterion validity assessment.

(Greg Stawicki): That was not specifically done as the testing of the measure was done in a mail survey format. But we did not send other measures to the respondents. And so we only have data really based on what they responded as part of survey and perhaps a little bit around their claims but not much other – in terms of other patient reported outcome measures that maybe part of what you're thinking here.

James Bost: Right. I mean, you know, your evidence table, I think there could have been some instruments that we're part of that. At least one potentially sent out with this instrument to assess, you know, correlation.

Robyn Nishimi: Anything else, committee members? OK, moving on then to threats to validity.

There were fair number of comments about whether there was a concern about language. Concerns were raised about the variable, geographic – whether it was medical complexity alone, geographic location alone or combination of the two for the risk adjustment. And then there was discussion about whether the developer did or should consider other variable. So with that, I'll open it up to the floor.

James Bost: I think that one was very well summarized.

Female: Yes.

Robyn Nishimi: Does the workgroup have any other variables they want to ask the developer about?

James Bost: I guess we could ask why they chose the one for risk adjustment that they did versus any others, or did they consider any others?

(Greg Stawicki): So we chose variables for risk adjustment that were readily available data from the actual survey instrument itself, so age and self-report health status comes out of the survey.

We did actually do some testing on gender and found that they had no impact. So we dropped it from a risk adjustment model. We did not have enough variability for race ethnicity testing to actually include it in a model. We didn't actually test it at that time.

This is two of our test in Medicaid populations. So, in terms of socioeconomic status, you don't have actually true SCS indicators or data but these are Medicaid population. So we did not include that and we certainly thought about other things.

Geographically, the data we have were from three distinct locations. Two state Medicaid plans and urban children hospital.

And so, again, with only limited number of sites, it really wasn't much else so we could put into a risk adjustment model without having appropriate end to do the appropriate risk adjustment.

James Bost: And is your opinion for norming and comparative assessments that risk adjustment is the best approach as opposed to stratifying by health status or age, and norming separately?

(Greg Stawicki): I think that it's an excellent question and I think when we did do the risk adjustment model, they really were very little differences in terms of actual scores. And likely it's because performance across these measures is still a quite low when it comes to the actual scores themselves. And we anticipate that adolescents, in general, are going to probably be answering questions pretty routinely. And we don't only expect to see a lot of differences.

Perhaps, we may by chronic condition type, and we certainly didn't have, you know, we certainly are not going to able to include, you know, 20 or 30 different conditions as risks (adjust) variable). So I think stratified analysis maybe a better way to go than risk adjustment. But for the two variables are pretty straightforward and they come out of the survey that was why we were able to develop that, that's a recommendation.

Lauren Agoratus: This is Lauren. I have a clarifying question on something you just stated. I'm wondering why you didn't consider SCS because of the Medicaid population only because that's still – Medicaid once a child was 18, parental income is no longer deemed. So the SCS would still be variable even though it's Medicaid population.

(Greg Stawicki): We did not have access to that data.

Lauren Agoratus: OK. OK.

Robyn Nishimi: Anything else on threats to validity? OK.

Feasibility, there were a lot of comments chiefly of concern was that it was mail only and telephone and electronic message aren't explored. So if the developer could comment on those.

(Greg Stawicki): So we chose to follow the development model and went through a mail survey. We acknowledge that electronic and telephone survey maybe an appropriate way to administer such as survey. There was a question around sort of – and with a telephone contract or an electronic contact because the target population was under age 18, getting consent from parents for this types of surveys at least in our setting. Was one of the issues also related the lack of being able to do it per se on telephone or electronic format. That's not to say that this couldn't be tested in the future using those kinds of formats.

Ricardo Quinonez: This is Ricardo, can I ask you something? How did the parents get – send the survey back you to, guys? Did you send free address, free stamp envelop?

(Greg Stawicki): Yes. Correct, yes.

Ricardo Quinonez: So what – so then – this was just retesting, how would you expect operationally to be that this would be administered?

(Greg Stawicki): I think that, you know, depending on whether it would be who would be doing the administration – from a health plan, or a clinical practice, or hospital. I think that, you know, that would be left to a decision of how that would be administered. I think similar to other patient experience measures in different places where sort of view survey vendors and other mechanism for survey return.

Ricardo Quinonez: OK. So this would incur some cost then through the …

(Greg Stawicki): Correct.

Ricardo Quinonez: … practice of health planning that's administering this. And then, the other common would be that people change addresses all the time but they rarely change e-mail addresses.

Robyn Nishimi: Anything else?

And the last item to discuss for ADAPT is the usability and use criterion. The comments were again a little bit disparate that we take home that I took away were one because the survey is new. No widespread use data are provided that that survey addresses the key concept but it doesn't address, again, whether the provider counseling is the best strategy. And then, yet, there was another comment that the measure is appropriate for accountability purposes.

So I think it would be useful for the committee work with members to discuss whether they think that the performance result will further the goal of high quality and efficient health care, that's the first question, and secondly, whether you feel that this is inappropriate measure for accountability purposes.

James Bost: So, this is Jim. I am very interested in hearing about the intervention, again, not being particularly amenable here. You know, I kind of figured that there must have been – I'm not a clinician, I'm kind of figured in the evidence and testing process that there were a lot of clinicians that agreed with this being an

appropriate way to assess. So I would certainly ask the developers to comment on that.

And I have to say, I appreciate and it really makes you think about usability if decision counseling is not an appropriate thing to assess or accountability and quality improvement.

Robyn Nishimi: Anyone else from the committee or the developer?

(Greg Stawicki): I think in terms of usability, I think there are certainly several avenues that this could be use in term of quality improvement and perhaps accountability related to, you know, whether it's an individual hospital, or practice, or a health plan, or health system. And I think that the point about, you know, whether or not this is actually reflective intervention and improve outcomes is a valid one to address but it's one that I think there needs to be starting point with discussion.

I think there is clinician – if clinician doesn't start with that kind of an approach. It hard to sort of say whether there's any type invention that's being done. That's said, you know, this is all well-taken comments around the usability of this kind of measure.

Robyn Nishimi: Anyone else? OK, great. We have walkthrough 2789 for the most part. I think the developers got some good feedback and the committee discuss where it concerns are and where they feel the strengths are.

So let's move on the 2770 which is the FECC measure. This is the Family Experiences with Coordination of Care, FECC.

So what I'd like to do here – this measure is a little bit different as you could tell, there were 10 measures within the submission form. And we reviewed it and, you know, had some specific questions for you to consider and then, you identified others. So the goal here again is discuss the differences of opinion which for this measure or for the 10 measures, there were many, on almost all the criteria.

So I'm going to try and walk us through those. At the same time from a logistic point if view, I will be periodically pressing you to weigh in on whether separate voting needs to be done on some of the 10 and it might be all of the 10. Because depending on reflection, different people called out different measures as being candidates for separate voting.

And the reason I'm going to press you is actually not a trivial matter for the staff to setup the voting in advance to the meeting. So we do want to have to the extend possible, relatively clear picture of what measures need to be voted on separate list, it might seem like a housekeeping matter but I'm just previewing for you that is actually not. So that's one going to stop you occasionally.

So moving then to the evidence, overall, there was a comment that the body of evidence for measures 1, 3, 5, 7, 8 and 9 with overall week with most of it relaying on RCT. For FECC 5, there was again concern about that five in that global comment. They would like the developer should address why the measure call for three-month period, is to be of three months when the evidence side monthly contact.

And then, I'm just going to – I'm going to take this up which are a little groups. And so for FECC 7, again, some clarification as to why they chose specific three-month timeframe, and I think that's mostly it and, again, a concern about the overall strength of body of evidence for FECC 7 relaying on one randomized control trial.

Is there anything else anyone would like to raise about FECC 1, 3, 5, 7. And I guess really, 8, because it's a global comment that's the evidence was weak.

James Bost: I think that your comment about three months definitely should be there – Jim here.

Female: Yes.

Robyn Nishimi: The developer on the line, it can be address why on the three months interval was chosen for FECC 5 and FECC 7.

Rita Mangione-Smith: Yes. So, this is Rita Mangione-Smith. So as FECC 5 which has to do with the frequency of contact between a designated care coordinator and the family of a child was medical complexity. It was in the study that was uses the evidence it was done monthly. The original quality – draft quality measure that we put in front of our (Delphi) panel did state the contact should happen monthly and it was the assessment of the (Delphi) panel in their in-person meeting here in discussion that that was stringent, it was the opinion of the Family Voices represented about in the (Delphi) panel that that would actually be an annoying level of contact. And that she felt most families would not want to be contacted that often, without them initiating contact themselves.

And so they debated and ask for us to revise the timeframe, and this was the measure they voted on so this is the measure we're approved for (inaudible).

First, FECC 7, which is the – one about the completion making sure that subspecialty appointments have been successfully completed by childhood medical complexity within a three-months timeframe. Again, asking that the care coordinator take on this responsibility. This, again, with a matter of very lively debate at our in-person (Delphi) panel meeting through these measures.

It was felt that the – it was reasonable to expect that within a three-month period when an initial referral is made by the child medical home to a subspecialist that appointment should have been made and occurred, and completed within a three-month timeframe and that if it hadn't – it was reasonable to expect that the care coordinator would intervene to expedite the appointment.

So, again, that was a timeframe that we put in there because none of the evidence suggest exactly what the timeframe should be. But in our own center workgroup that worked on these measures, it was decided this was reasonable in our (Delphi) panel through their discussions endorse that they felt that the three-month time period was reasonable for seeing completion of a subspecialty visit.

Ricardo Quinonez: Rita, can you tell us – this is Ricardo. Can you tell us what outcome was that – was tested in the RCT that supports FECC 7?

Rita Mangione-Smith: So, Ricardo, this is a bundled intervention, so it's very difficult for us to call out that any of the outcomes that were looked at in the – the specific RCT would have been the explanatory, you know, that they could be explain just by FECC 7. So that this was part of the weakness of the evidence that we face, nobody has really pulled this various pieces of the evidence apart and look at them individually in terms of there impact outcomes. That particular intervention, I'm going to have to look back at it – Go ahead.

(Off-mike)

Female: I would need to look at that actual table of outcomes.

Rita Mangione-Smith: OK. So (KC Lyon) is here with me and she was very involve in the evidence synthesis for this. So we're just looking at their outcomes.

(Off-mike)

Rita Mangione-Smith: OK

(KC Lyon): That was the (Janette Farmers) 2011 paper which was the randomized controlled trial crossover to intervention, and they've found and that included. So it was care coordination and they have a lot of that (elements) that ended up factoring into our various measures as part of their intervention and they found to increase satisfaction, decreased family stream, decreased unmet needs.

Female: I think it goes further over each, sorry. Oh, we got cut off.

(Off-mike)

(KC Lyon): So they ended up looking at a number of different outcomes. But, again, as Rita mentioned, it was a bundle of intervention, was the most important so although we know that bundled interventions are more likely to be effective than individual interventions on their own, so perhaps it's hard, right? We

probably couldn't tease out what the effective any individual piece was at any rate.

Ricardo Quinonez: This is Ricardo again, sorry. These metrics in, you know, taken together or individually, they are meant for a health plan mostly?

Rita Mangione-Smith: That is the level that which they were tested Ricardo. So that is the level at which we are suggesting they should be use.

Ricardo Quinonez: And did you survey any health plans to see what the possibility of having a care coordinator for every child with complex care needs would be since it is a big part of this?

Rita Mangione-Smith: And so we did not survey health plans regarding that question, no.

James Bost: So, this is Jim. What do you consider the denominator for these measures?

Rita Mangione-Smith: So the denominator was arrived at using the pediatric medical complexity algorithm and required that the child be categorized by the algorithm using it three – two years worth – I'm sorry, three years worth of administrative claims data and it's like complex chronic. Complex chronic means you have either two active body systems involve or that you have one active body system a condition that is considered to be life-shortening. So an example of that would be cystic fibrosis. Or, that you have a malignancy that hasn't been a remission for at least five years.

So that's how the algorithm functions to pick out who is complex chronic. In addition to that, we require that children had four visits to their health care provider in the prior year in order to tag children who are using lots of health services and very likely are more likely to need care coordination.

So we really were trying to find the denominator population that would benefit most from care coordination. And, I guess, in reference to your question, Ricardo, this would actually be a relatively small percentage of any health plans population but certainly the percentage that would have very high utilization and extremely high cost.

Ricardo Quinonez: And there's a thought that having a, probably, probably a caregiver would be decrease this cost.

Rita Mangione-Smith: The care coordinator?

Ricardo Quinonez: The care coordinator would probably decrease this cost, right, since you're providing better quality.

Rita Mangione-Smith: So there is mixed evidence regarding that in the literature.

Ricardo Quinonez: OK.

Rita Mangione-Smith: We would hypothesize that having better care coordination would prevent things like inpatient hospitalizations that were not of an acute nature, so unplanned admissions and as well as E.D. utilization. There would be an increase, however, in the use of outpatient services …

Ricardo Quinonez: Right.

Rita Mangione-Smith: … we've seen that in many of these studies that outpatient cost go up while inpatient and E.D. cost go down.

Ricardo Quinonez: Which are this proportionately higher so that …

Rita Mangione-Smith: Right.

Ricardo Quinonez: … makes a lot of sense.

Rita Mangione-Smith: Right.

Ricardo Quinonez: Do you worry at all that this would limit the number of health care plans that this population would have access too given that it would represent significant expense to at least a percent of their patients for the health care plans?

Rita Mangione-Smith: Am I concern that health plans will carve this children out, is that your question, Ricardo?

Ricardo Quinonez: Yes, and that is population would have of less choices for health care plan.

Rita Mangione-Smith: Yes, I am certainly concern about that. I think as our most people who take care of these children.

Ricardo Quinonez: OK.

James Bost: Yes, so this Jim. To get back to why I ask the question, we are in the process here at Children's Healthcare of Atlanta, I'm on the design team for developing an inpatient complex care center and an outpatient complex care center where the real focus is around coordinate of care and trying to get all your specialist together in one center. And I think that we would probably have issues with the three-month designation for 7 being a little far out there for complex kits.

Rita Mangione-Smith: Yes. So I think we heard both opinions of the (Delphi) panel. There were people who felt it should be much more stringent and there were people who work within systems that do not have that kind of multidisciplinary access within their own roof or within their own system. And felt that if you're having to refer outside that it was too stringent to make it a months which was the original suggestion. But you would see other children where we have that kind of everything is under one roof thing, you know, I agree with you, three months is a lot.

Robyn Nishimi: So let me ask the workgroup, did they want to split out voting for 5 and 7 because they both have the three months – that's part of the specification.

James Bost: This is Jim. I don't think so at this point, just based on that criteria but that just my opinion.

Robyn Nishimi: Anyone else? OK.

Let's move on to – I'm going to take up FECC 9, 14, and 15. For 9, which is caregivers receiving written summary. They were actually a few comments. There was one comment that thought that this was pretty standard for written summary. There were two comments about they were missing components list of hospitalizations, behavioral plan is applicable, letters to the emergency room providers is applicable and all allergies, not just medication allergies

should be included, so the developer might wish to address that. And then more – and then just generally, the comment that more information was needed regarding, you know, what exactly the written summary meant sort of, I guess, subcomponents if you will.

(KC Lyon): And so, this is (KC Lyon). And the majority of FECC 9 is drawn from that NCQA standards and guidelines which was one of the documents that actually listed out specific elements that should be included in the summary. In part of the guidelines, they actually say only that the patient or family should receive a clinical summary at each relevant visit. But later, it's falls out that when appropriate diagnosis medication is recommended treatment and follow-up should also be included. So that's why a part of the specification for what should actually be a part of the summary came from.

There were additional elements such as allergies so specialist involve and what to do for problems related to the outpatient visit that really came from that discussions with the (Delphi) panel, again, based on face validity.

So some of the same comments that it sound like came at the initial review that some people felt like the standards laid out of just diagnosis medications recommended treatment follow was not adequate and that there should be more included.

And so the (Delphi) panel recommended strongly that we include those additional allergy specialists and what to do for problems related to the visit. And so that was the measure that they develop that they voted on. And so therefore what we are presenting but this was another topic of active discussion about what specifically should be required.

Ricardo Quinonez: This is Ricardo Quinonez. While there is not much evidence right now about how discharge summaries impact subsequent use of health facilities, readmissions, et cetera. A lot more is coming on that and I think having a little bit of insider's knowledge with some of the groups that are doing this, is absolutely going to be pen out that that having a standardize discharge summary will actually prevent re-hospitalizations, going to the emergency room particularly with the kids with special health care needs.

The section on hospital medicine at the AAP through it's quality subcommittee is doing a multicenter quality improvement project that includes outpatient provider to sort of standardize an instrument for discharge summaries that I think it has in his initial phase shown that the standardized is well-received by the pediatricians in the outpatient setting and that it decreases use of E.R., readmissions, et cetera. So I think a lot more is coming to support the validity of this recommendation.

Robyn Nishimi: Right. So I guess the question would be, Ricardo, what's specified here is appropriate to whatever the standardized instrument is. Not about the importance of a discharge summary but whether the measure that specify needs the same …

Ricardo Quinonez: I think what the evidence is going to show is that the – and is important to use a standardized instrument.

Robyn Nishimi: OK.

Ricardo Quinonez: I don't know that they're testing each particular part of the – standardized instrument. Just that it is much more beneficial for providers and hospitals who see a lot of this patient in the inpatients setting that it would be a standardized instrument.

Robyn Nishimi: OK. And then, in this case, it has been standardized. Are there anymore questions or comments by the – for the committee? OK.

Let's move onto FECC 14. This is the communication with school staff. There were (inaudible) opposed comments from workgroup members on this. One felt that studies in most chronic disease population – so that interaction with the school improves outcomes. And then another comment was that the evidence for measure 14 is barely tangential. Involve one trial with 56 children with TBI. And so was – insufficient to justify a broad measure asking for communication with the school, so committee want to discuss this.

Female: Is that to us or to the panel?

Robyn Nishimi: But I'm asking the committee to discuss their opposing opinions.

James Bost: Yes. This is Jim. I think the way you said it is right. I mean, the – what was presented was not a lot of evidence. However, I've work in a lot of areas and not necessarily be the way the denominator is chosen here but other children with chronic conditions and have seen studies where such interaction with schools, to improve care. So I guess I was thinking about this, just in terms of other evidence I've seen.

But, you know, as you've said, the evidence provided was not very strong for this population.

Robyn Nishimi: Any other comments? So then, that would be something to raise with the larger group, Jim, that your perspective on other evidence.

Kerri Fei: This is Kerri Fei as well. I had made a comment to that effect as well, wondering if there was other evidence from other chronic disease populations that could be found to be applicable to support that measure.

Lauren Agoratus: This is Lauren, I had the same concern.

Robyn Nishimi: OK. So then maybe the developer can come prepare to address that at the meeting.

Female: OK.

Robyn Nishimi: Moving on to 15, which is professional interpretation for non-English speaking patients. Again, there was a little bit of concern that the evidence for 15 was – not graded. But that, in general, there was another comment that evidence for measure 15 and also 16 that we're talking about 15 was acceptable. And then another that comments at professional interpretation really is important. Is there anything that committee feels they need to discuss about the evidence for 15, which is on interpretation? OK.

The last two for 16 and 17, the comments around 16 and it went here as well as when we went to the validity of the specifications was the notion of who the main provider was. So that was of concern and the committee might want to discuss that with the developer.

And then for 17, there was concern that the – what was offered was a consensus statement not a systematic review. There was no empirical evidence. And there was actually recommendation to separately vote on 17 if the AAP was, in fact, just based on consensus opinion.

So let's talk first about 16 and the issue of main provider who develop. Who want to address that for the committee?

(KC Lyon): Sure. This is (KC Lyon). So the – OK. So Rita's point, the definition that we use in the survey itself, but the idea was that when were initially formulating this, we were very much thinking about the primary care provider within that Patient Centered Medical Home.

But something that came up in a lot of the discussions that we had was that, some patients actually have their medical home in a subspecialty clinic. So for instance, the patient with cystic fibrosis may actually identify their pulmonologist as their main provider within their medical home provided within the pulmonary clinic. And some of that is going to be very system dependent.

And so we wanted to have a definition that was flexible enough to incorporate all of the various systems that I use in different places, and not unfairly penalize. Someone who is assigned as a PCP, who actually doesn't function as the child's medical home.

So the definition that we use in the survey is, we're started out the survey – which your child's main provider is the doctor-physician, assistant nurse or other healthcare provider who knows the most about your child's health and who's in-charge of your child care overall.

And which we deal with the providers, please tell me what should these providers as your child's main provider. And we then provided a list of providers that the child had seen in the previous year for them to select. And that would be the level – that would essentially who we would attribute many of these quality measures too.

Robyn Nishimi: OK. Does the committee have any questions about main provider given the response from (KC)?

James Bost: This is Jim. I'd like to think that, when I've read your evidence, it seems like a lot of these we're being done by care coordinator or potentially case manager that doesn't sounds to me a like a main provider.

Rita Mangione-Smith: So the questions in that survey are phrased such that we'd ask about whether there's anybody in the main providers and services, and if they answer, yes, we ask them, you know, who that person is. And it can be – they can check that is their main provider, that it's another doctor or nurse in the main provider's office. That if a care coordinator, a social worker, a care manager, some else in the main provider's office.

And then the questions are phrased such that we talk about the person in the main provider's office who helped you with managing your child's care. That phrase is used in every survey question. So it is not suggesting that it's the main provider who would be the care coordinator.

James Bost: And so what you describe that's really tells me that you can only to do this by phone or in-person.

Rita Mangione-Smith: We did it by mail and by …

James Bost: I'm not sure how you've got all of that that you've just said in a mail survey.

Rita Mangione-Smith: We did. Happy to show you the write up.

James Bost: Yes. I mean I – through it, just I think if I was reading, I don't know that I would take the time to look at all that when answering the question, but it's there.

Robyn Nishimi: OK. So that goes to the issue of reliability. And we'll get to that in the modes of administration.

For FECC 17 then, again, the concern was that the – it was the consensus recommendation by the American Academy of Pediatrics. No other empirical evidence, and there was a recommendation to – for the separate vote on 17.

Can the developers confirm that this was just a consensus expert opinion-based?

Rita Mangione-Smith: Yes.

Robyn Nishimi: OK. So I'm given this, I'm going to ask if there are any objections to a separate vote on FECC 17?

Ricardo Quinonez: Well, can the developer – this was a consensus statement from the AAP but also tested through a (Delphi) panel, correct Rita?

Rita Mangione-Smith: All of these measures were assess for content validity and feasibility of implementation by a 9 member multi-stakeholders (Delphi) panel yes.

James Bost: OK. I think we should take that into account when looking at the evidence for this.

Robyn Nishimi: Well, but we're getting to – we get that to that under validity. Right now, we're in the evidence section.

Ricardo Quinonez: OK.

James Bost: I guess my question would be, it is from a consensus statement from the AAP but when they we're developing that, did they do evidence-based research that they could share with you to add to this?

(KC Lyon): So the consensus statement itself – this is (KC Lyon). The consensus statement itself specifically says while there is no empirical evidence to support this. But if they did mention in the statement that they know – have multiple – that the committee was aware of multiple locations where these are being implemented. I supposed we can reach out to them to see if there's any data to support that. But at least as of the consensus statement there, they were aware if people using them and being happy with them but did not present any data.

Lauren Agoratus: This is Lauren. Can we get the citations for that consensus statement?

Kerri Fei: Yes. This is Kerri. I was going to ask the same. I went looking for it and I couldn't find it. That's why I'm asking.

Lauren Agoratus: Thank you.

(KC Lyon): Sure. Hold on. We're on – we can – if there are way we can send that around …

Robyn Nishimi: If you just send it to the feeds NQF box, we'll take care of distributing that.

(KC Lyon): OK, great.

Rita Mangione-Smith: We can do that. We'll be back.

Robyn Nishimi: So right now, we would be going into the meeting with a separate vote pending perhaps empirical evidence. But right now, it sounds like the recommendation for a separate vote on 17 should occur.

James Bost: And then it's because it would pull the rest of them down because there's no empirical evidence?

Robyn Nishimi: Yes. That's what I've heard on the call. And if you follow the algorithm, yes. that will pull it down.

James Bost: Yes, the algorithm.

Robyn Nishimi: OK. I'm a little bit concern about our time. So let's move on to gap in care.

Overall, there was general support that there was a gap in care. It was viewed as a low occurrence high-risk population. So that made measures valuable. It was noted though that no performance data were actually supplied. So there was a concern that you couldn't tell whether there was a gap or not.

Any thoughts on the – whether there's gap in care or whether there's insufficient data provided to assess this? Any thoughts? OK.

Well then we'll just let the comment stand, some people thought there is a gap and some people thought there was insufficient information.

Moving on then to the reliability of those specifications, for the most part, the – and specially now that we've cleared up the issue of the denominator and the main provider construct. The sense of the workgroup comments was that measure – or the measures rather could be consistently implemented.

There was concern that given the developers state of these measures can be use independently of each other. Each one should be voted on separately.

Are there any other comments about the clarity of the specifications, the algorithm numerator, denominator? OK.

Then let's go to the reliability testing that this is really where we will get into more on whether we need to split this out further.

Concern with express about the result for FECC three and FECC 15 due to small sample sizes, there are multiple comments about FECC 3 and FECC 15 in this regard, and the fact that for 3 and 15, they couldn't even report result so developers couldn't.

There was also concern about FECC 5 FECC 8, about the reliability of that. Just because there was internal consistency in the output for them, the question was whether really a test-retest for reliability should have or could have been done that would approving – better reliability for 5 and 8 rather than just internal consistency.

So the concerns, again, are really around 3, 15, 5 and 8. Does anyone else have any questions about other reliability testings for the others?

So if anyone objects to splitting out 3, 5, 8 and 15 given the several issues raised by the workgroup?

Just to remind you, 3 is caregivers of children with medical complexities who report having a care coordinator and that the coordinator has contacted them in the last three months. So that's, I'm sorry, that's 5. 3, is report having a designated care coordinator. 8 is report having a care coordinator should also report that their care coordinator was knowledgeable, supported the caregiver and advocate for the needs of the child. And 15 is the language and

professional interpretation, so those would be the one that we would pull out separately.

Are there any others concerns about the reliability testing?

James Bost: No. I think you did a highlight the ones that are potentially should be scored independently. I guess, I would ask the developers at this point their thoughts on the fact that we are choosing to potentially score this separately as a different way to endorse them.

Robyn Nishimi: Well, the developers say that they are independent measures.

James Bost: Right. So, they are willing to have us do them separately as an endorsed – as individual items for endorsement?

Rita Mangione-Smith: Absolutely. Well, you know, the original, the entire survey actually encompasses 20 measures and we internally chose not to put 10 of them forward because of various – it's under weaknesses in them in terms of evidence reliability and validity. These were the 10 strongest out of the 20. But, you know, we certainly understand that they are not all equally strong and we do see them as individual measures. So, we don't have a problem with that.

Robyn Nishimi: OK. So, we will move forward then with 3, 15, 5 and 8, and in addition 17 because we had already pull that out underneath the evidence section. For validity, again, there was concern about the underlying evidence for 5, 7 and 9, and the – we've already discussed the evidence for 14, the overall folks felt that the rest of the evidence – the rest of the specifications more consistent with the evidence that was presented.

So, 7 and 9 because we have not address those separately, that's being pulled out. 7 is caregivers who report, having the care coordinator should also report that the care coordinator assist them with specialty service referrals by ensuring that the appointment with the specialty provider occurs, so that's 7. And 9 is the written discharge summary.

So, I think, we've discussed the evidence for 9 fairly well, but do folks have concerns about the underlying evidence for five which then would affect whether the specifications are valid. This is the three months one.

James Bost: And 15 had, I think, low convergent validity but we're already doing that one separate so.

Robyn Nishimi: Right. OK. So, doesn't sound like we need to pull anything out based on the validity of the specs. Let's go to the validity testing.

Again, a dichotomy of opinion validity testing is week versus the comment of validity testing was comprehensive. So the workgroup should discuss these varying opinions. See if they could bring a little bit more clarity to the issue for their follow-up committee members.

James Bost: I guess – go ahead. Sorry.

Lauren Agoratus: I'm sorry. This is Lauren. My concern mainly in this area was number 15.

James Bost: Yes.

Lauren Agoratus: So, you know, small sample size and so forth.

Robyn Nishimi: Jim, you were going to …

James Bost: Yes. I was just going to say, I think that, you know, the approach they did, it cover a lot of areas of validity. I think they covered content (phase) and convergent validity fairly well. The fact that 15 did not correlate well with the other items, did not surprise me because I think the medical interpretive question is the least similar to some of the other things they we're looking at, but you know, it did.

So I think that, unfortunately, that wasn't addressed well for 15.

Robyn Nishimi: OK, anything else on the validity testing? OK. Let's move on to threats into validity.

There were several comments here about the threats to validity. There was concerned about 17. But again, that went to the issue of evidence and so we've already address that. Really, most of the comments express concern about addressing for survey mode, the education level issue and whether the lack of demographic data will adversely affect the validity of all the measures, and the rationale, again, for why education and mode of administration for risk adjustment were put forth.

So I think, there are two things I'd like the workgroup to discuss first, or any additional variables and the lack of demographic data that impacted that on the validity of the measure overall, measures – all 10 of them overall. And then we could then ask the developer about the adjustment for survey mode and education.

Are folks now not concerned about the lack of demographic data and the validity?

James Bost: This is Jim. I didn't have particular issues there. But not a clinician, again, I don't know whether education was the only choice. And as you said, we'll hear from the developers what other variables they considered for risk adjustment.

I also like the developers to comment on the need for risk adjustments for some of these models. When I look at the coefficients, they were kind of small. And so, you know, I'd like to know whether they believe risk adjustment should be considered by individual question again as opposed to overall.

Robyn Nishimi: Anyone else? OK.

The developer, if you can respond to the questions related to adjustment for the survey mode and the education and what other variables you may or may not have considered.

(KC Lyon): Sure. This is (KC). So, we consider – so the reason we need adjust for education and in spite of the fact that some of the adjustment coefficients were fairly small. It's largely based on our experience with survey response

patterns based on specifically caregiver education, which is commonly seen in survey such as CAHPS. Where we see that there's actually differential response patterns.

So people interact with the survey response options in a different way based on the level of caregiver education, so not necessarily reflecting differences in the quality of care received but really differences in a way, and essentially the likelihood of choosing the most extreme response option versus using some of the more, some of the response options in the middle of scales. So that was largely driven by theoretical concerns and was more of an (A) priority decision.

With regard to the survey mode, we didn't see differences in response patterns by survey mode and so that's why if you were using multiple different survey modes, we would recommend adjusting for that. However, if you chose only a single approach, you clearly that would obviate the need for that.

And then with regard to other demographic characteristics that we considered adjustments for such as things – race, ethnicity, language, we're concerned – gender …

Female: Age, all of those.

(KC Lyon): Age – so things like gender and age didn't seem to make a significant difference and we didn't have strong reasons. Caregiver gender would have been very hard to adjust for because the majority of our respondents were female. And we didn't – with regards to things like race, ethnicity and language, we are concerned that differences that we might find on the basis of those characteristics represents true differences in care being received so we would not want to adjust delay, right. That would be a case where, perhaps, want to stratify to look for disparities. But wouldn't actually want to adjust the way the differences in care that are being received.

James Bost: The education categories, did you choose them from some source?

(KC Lyon): We used the same categories that are used on the demographic questions for the CAHPS instrument.

James Bost: Yes.

Robyn Nishimi: Anything else related to threats to validity?

James Bost: So you said that the majority were female, yes, what was the ratio?

Rita Mangione-Smith: It was about 90 percent of female caregivers respondents which is very common in any pediatric study were there's proxy reporting.

James Bost: Yes. So you don't know whether gender wouldn't matter but like you said, you know, you had, I don't want to use the term bias but you currently had a, you couldn't assess that because of the large female respondents.

Rita Mangione-Smith: (Inaudible) we don't have any reasons to believe it would be any different if this survey was given to another set of parents more widely. And again, that's based on experience with child CAHPS measure, so there – uniformly majority of respondents will be mothers or female caregivers.

Robyn Nishimi: Anything else? OK. Let's move on to feasibility. The comments for feasibility were a question was raised whether in person could be an option? The administration generally seemed feasible. Since the developer recommends the mix mode approach, should there just be a single recommended mode of administration in which case you wouldn't need to risk adjust. And then there was also a comment that overall – the measure – survey measures in general can be very cumbersome and so there was a concern about the feasibility of the measure on broad national scale.

Any other points the committee wants to make?

James Bost: So that issue about burden, I guess, that comes back to – if these are items that are really individual in nature, is the feeling that you can pick and choose which one do you want to use as oppose to doing them in set. So, it's hard to know burden because maybe you will only pick two or three.

Robyn Nishimi: Correct. (KC) or Rita, can you address the recommendation for a mix mode and whether you should just choose one mode of administration.

Rita Mangione-Smith: Yes. So, what we found in our mode analysis was that if you were to go strictly with a mailed survey mode, thus part of the population we would lose would be those who are in minority population and those who are non-English speakers. When we went with the mixed mode, we had the highest response rate and we were able to then capture many more low English proficiency participants and many more minority race, ethnicity participants. Mailed surveys were largely returned by non-Hispanic white English speakers.

So for cost reasons, we suggest a mixed mode because mail is clearly cheaper than telephone interview, and if you can get a reasonable response rate from non-Hispanic whites by mail, you don't have to call them.

But the mixed mode, if you don't get a response initially from somebody, the chance that you will capture a more representative population of caregivers goes up if you use the mix mode.

And I also just wanted to clarify one thing. Mixed mode, you would still have to risk adjust for whether or not you got a mailed survey back or a telephone survey back because we did see variation in scores depending on whether you responded by phone or by mail.

(KC Lyon): OK. I was …

Robyn Nishimi: Thanks.

(KC Lyon): This is (KC), the other thing I would add is that in the telephone only group, we actually saw much lower rates of participation in the non-Hispanic white and highly educated population. So if you went with – for simplicity's sake, if you want an adjusted health on population, you would also be missing part of the population you're ideally interested in capturing.

James Bost: OK. A couple of questions, the (found) administration, was that only if they had a home phone or did you also consider cellphone or mobile phone?

Rita Mangione-Smith: We could only consider the phone number information we were given by the Medicaid agencies. And that may have been cellphones, that may have been landlines. We have no way of knowing that.

James Bost: So the – you said – you talked about the significance of the mode of administration but I guess what I'm looking across is the P values, it looks like only FECC 16 was significant.

Rita Mangione-Smith: I don't – I'm not sure exactly where you're looking but I'm pretty confident it was more than just one measure where we saw mode effect.

James Bost: Yes. I was looking at the detailed risk models back to the Excel spreadsheet we were given.

(KC Lyon): You know, so I will, Ill have to look back at says, full. At the – our full result – there may have been more of the differences, may have been in some of the measures that we didn't submit.

Rita Mangione-Smith: So that's right.

(KC Lyon): For NQF endorsement, right? So …

Rita Mangione-Smith: Yes, yes.

(KC Lyon): So this is sort of half of the overall subset.

Rita Mangione-Smith: Right, because this is just 10 of 20 measures that are included in the survey, which I'm forgetting though, yes.

Robyn Nishimi: OK, anymore discussion around the feasibility? Then let's move to the last criterion to discuss today for the FECC measures. The comments on usability were on the – mail issues with the usability of it, that the information gained through the tool could be very important and that the benefits of the tool and the measure could outweigh the potential consequences.

Any other comments that people wish to make on usability? OK.

So we've marched through, I think, we've identified the ones that we want to pull out and then the ones that can be grouped together. Are there any other comments the committee wants to make in general about any of the 10 measures before I turn it back over to Suzanne?

James Bost: I guess, I would ask the question, do we want to just do them all separately at this point?

Robyn Nishimi: Sorry that's …

James Bost: … since we pulled that pretty much more than half I think.

Marlene Miller: This is Marlene. I think that will probably make sense.

Kerri Fei: This is Kerri …

Robyn Nishimi: Does anyone object?

Kerri Fei: Sorry. This is Kerri. I actually had the same thought as well, about pulling them all out separate.

Robyn Nishimi: But let me put it this way, does anyone object to making them all separate?

Amy Houtrow: This is Amy, I do not object.

Robyn Nishimi: OK. That makes our housekeeping much easier. Anything else? Suzanne Theberge, turning it back to you.

Suzanne Theberge: Thank you, everybody. OK. So I just want to go over the next – sorry, public comment and then we'll do our next step.

So, operator, we'd like to open the line for public comments.

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

Suzanne Theberge: And please feel free to submit a comment via chat as well.

Operator: And there are no public comments at this time.

Suzanne Theberge: OK. I don't see anything on the chat either. OK.

So, we'll move on to next steps, next slide please.

Committee, your next steps are to review the remaining measures in the project. You've now looked at least the FECC and the ADAPT measures and we'd now like you to start looking at the remaining measures. And we ask you to review everything prior to the in-person meeting. You don't have to fill out the surveys anymore unless you would like to. If you find that helpful please feel free to just send it, you're definitely not obligated.

And we will be getting in touch with you no later than early next week, hopefully sooner, regarding your lead discussant role. And basically, the lead discussant is we asked a couple of folks to be responsible for leading the discussion on each of the measures, you know being kind of team to kick off the discussion, raise the issue that were discussed on the workgroup call. So we'll be getting in touch of these hearing about which measure or measures you're assigned to. So keep a look out for that e-mail.

And also just to – on kind of logistics perspective. If you have not already does so, please do make your travel arrangement for your trip to D.C. and we're looking forward to seeing you soon. You should have received information about that from our meetings team, so if you didn't, please let us know and will get that fixed.

Also, I just wanted to know we'll be sharing the rest of the workgroup comments with you and the public comments on the measures in the next couple of weeks.

Next slide, please. I'll just go over the remaining calls and that, next to workers or next week Monday and Wednesday. You are welcome to listen in. But again, you don't have to but please feel free to go in or stream the audio if you're interested in the discussion. And once the workgroups are completed and the public comment period that we're having before the meeting is completed, we'll be updating the measure worksheet with the comments that

were submitted by you and the members of the public. So that will be in there as you review the measures.

And as I said the in-person meeting, December 1st and 2nd, and then if we don't get do everything at the in-person meeting, we do have a follow-up call scheduled for December 10th for 3:00 to 5:00. So what we should know by December 2nd whether or not we need that call.

I will pause here and see if there are any questions.

All right, next slide please.

The next slide has our contact information. If you do have questions on process on the criteria on logistics anything, please don't hesitate to e-mail the team or give us a call and committee members please do check the SharePoint page for updates, so we'll be e-mailing you as well as we post more information or updated versions for the worksheets with workgroup comments, et cetera.

So I think, unless anybody has any other questions or comments, that concludes our call for the day.

And, again, thank you so much to everybody who took a couple of hours to participate. We really appreciate it. And now, enjoy the rest of your day, have a good day and the rest of your week.

James Bost: And thank you, Robyn, for leading the discussion today.

Robyn Nishimi: My pleasure.

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

END