NATIONAL QUALITY FORUM

Moderator: Taroon Amin
March 19, 2013
1:00 p.m. ET

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

Suzanne Theberge: Good afternoon everybody, and welcome to the GI/GU Stage Two, Steering Committee Orientation Call. Thanks for joining us today. We are excited to get started on this next phase of the project and we're looking forward to working with you again.

This is Suzanne Theberge, I'll be managing the project this phase and the rest of the team will introduce themselves in a minute after we go through the Steering Committee.

So, I'd like to just – sorry, I guess we have a different slide set than everybody else here. The project team is myself and Ashlie will let you introduce yourself.

Ashlie Wilbon: Hi everyone, this is Ashlie Wilbon. I think I've met everyone on the call by now in Stage One. And I'm working with Suzanne on this stage as well to make sure that we can get through direct lead measures. And looking forward to working with you guys again.

Suzanne Theberge: We also have Evan, here.

Evan Williamson: Yes, Evan Williamson. Excited to work with you guys again on this stage of the project.

Suzanne Theberge: All right and we've had a transition for this stage, we'll have – Reva Winkler will be our Senior Director on this stage. Taroon will be stepping off. And I'll let Reva introduce yourself.

Reva Winkler: Hi everybody, thank you all for working with NQF. I'm Reva Winkler, I'm a Senior Director here at Performance Measures, and I'm just beginning my 13th year of working at NQF with endorsement of measures. So thanks all for being with us.

Suzanne Theberge: OK. So now we'd like to go through the Steering Committee and we'll just do a rundown of which developers are on the call, as well. I am going to read off the list of the Steering Committee members, because we're not – we know we've got some people dialing on Friday instead of today.

And we're also going to ask you to do a brief conflict of interest review when you do your introduction. So, what we'd like to do is we'll go through the list and then have you say that you're here and then make any disclosure that you wish to make. Just because you wish to make a disclosure, does not mean you have a conflict of interest. It's just letting us know the thing that might be relevant to your work on this project.

We don't need to go through you C.V. We just want you to know about topics that we'll be discussing today on this call particularly consulting, speaking engagements, grant money, research money – anything that would be relevant to the measures in front of us today.

We also want to make sure that you remember conflict of interest and disclosure is not just financial. A lot of committee members will say, "I have no financial conflict of interest." And that certainly could be part of this scenario. But because many folks that are on our committee do volunteer work as well, we would like to make you – make sure you disclose that as well. Just anything that may be related to the measures.

And finally, we just want to make sure you all remember that you're serving as individuals. You're not here as representatives of your employer or of the organization that nominated you, or anything like that. You know, you're not

here as a – representing the American Society of something. You're here because you're an expert in and of yourself.

So with that, I will run through the list starting with the Co-Chairs. Andy Baskin is one of our Co-Chairs to help me on the Friday call. Chris?

Christopher Saigal: Yes, I'm Christopher Saigal, I'm a Urologist at UCLA. My disclosures are I'm a co-founder of a company called WiserCare, I think it's related to this. I'm also (NIHPI) on topics that do relate to, you know, GU or quality of care.

Suzanne Theberge: Thanks. Liliana Bordeianou? I think she's calling Friday. Zahid Butt?

Zahid Butt: Yes, hi. I'm a practicing Gastroenterologist in Columbia, Maryland. Member of a 16-person GI group. I'm also a CEO of Medisolv that provides implementation of electronic quality measures for hospitals and health systems. I do not have any specific disclosures for this project and I have no conflict of interest.

Suzanne Theberge: Thanks. Robert Ellis?

Robert Ellis: Hi, I'm a Director at the Center for the Study of Services and Consumers' CHECKBOOK. And I have no disclosures to make.

Suzanne Theberge: Nancy Faller? OK. Edward Gill? OK. Johannes Koch? Jenifer Lightdale? She emailed and said she'll be dialing in late, so, she may not have joined us yet. Richard Luetkemeyer? I think she's also coming in Friday. Alayne Markland?

Alayne Markland: Hi, I'm Alayne and I have – I'm currently On-Staff Associate Professor at UAB as a Geriatrician in the Department of Medicine. I also work for the Department of Veterans Affairs. And I do receive a research grant funding from the Department of Veterans Affairs, NIH on the clinical treatment of (inaudible).

Suzanne Theberge: OK, thank you. Paul Merguerian? OK. John Morton? Anne Pelletier-Cameron?

Anne Pelletier-Cameron: I'm on the call. I am Pelletier-Cameron, I'm a urologist at the University of Michigan. For closures I'm a co-investigator in an NIH funded a study on Bowel and Bladder Management after a spinal cord injury. And the PI on a study on post prostatectomy and continence although I received no financial remuneration for that.

Suzanne Theberge: OK, thank you. Stuart Reynolds?

Stuart Reynolds: Yes. Hi, this is Stu Reynolds. I'm a urologist from Vanderbilt University and I have no disclosures to declare.

Suzanne Theberge: OK, and Philip Schoenfeld? OK.

So, I know there's a number of developers who have dialed in. If you can just, real quick, say your name, and what organization you are with, that would be great.

Erin Giovannetti: This is Erin Giovannetti, I'm with the National Committee for Quality Assurance.

Suzanne Theberge: OK.

Female: Hi, Eden.

Eden Essex: This is Eden Essex with the American Society for Gastrointestinal Endoscopy.

Female: (Inaudible) colleagues with from AMA-PCPI.

(Bonnie Veer): (Bonnie Veer), I am colleagues with ActiveHealth Management.

(Judy Flug): (Judy Flug), Quality Manager, also from ActiveHealth Management.

George Wu: George Wu, from ActiveHealth Management.

Bonnie Paris: Bonnie Paris from Quality Quest.

Suzanne Theberge: All right, do we have any other developers?

OK. Great. Committee folks, just so you know the developers are here as your reference if you have any questions as we go through the checklist. You can ask those questions.

So, OK. Just to quickly go over the agenda for the call. We are going to go over the Stage One Checklist and talk about whether the developers answered your questions on Stage One. And in that discussion, which will be facilitated by Chris and Ashlie, we will be deciding whether the measures that are going to be discussed in Stage Two.

And then, once we're done with that, we will give you an overview of measurement. Talk about NQFs Measure Evaluation Process for the remaining three criteria; and then answer your questions.

All right. So now I'm going to turn it over to Evans to talk about the CDP.

Evan Williamson: Great. Before we start, I want to make sure that everybody sees the slides on their webinar – everything's coming through properly.

Male: Yes.

Evan Williamson: Great. Excellent. All right. So, moving on to the Business Development Process. When we started this project we did a two-stage process and we went through the First Phase, as you all know, to focus on the importance and to measure and report to redo the concept of the measures.

And here in Stage Two, we're going to do the fully specified measure – I want to focus on the scientific acceptability, the feasibility, and usability, the meat of the measure testing. This process will go through a similar process to Stage One that we went through. We'll review it at the Steering Committee. We'll go through an approval process. We'll go out for another comment period, in addition to the member comment that's already happened. Then we'll go for a vote, then we'll go CSAC again, then finally for board ratification where it will become an endorsed measure.

And here we've laid about in sequential order. You can see, these measures went through technical assistance after they were approved as a concept. Not

all the measure developers have to submit at least one of their measures to go through the technical assistance process.

You can see that the developer must "demonstrate the checklist for Stage Two review" has been met. So, during the First Stage, we've put together a checklist recommendations that we'll be going through today on this call. And the developer had to – had to justify the checklist and – during their submission for Stage Two.

We're going to have conference calls to discuss these measures during the Steering Committee Review. And (inaudible) comments, member voting, and then finally CSAC and board approval.

Do you have any questions on what I've covered, so far?

Male: No.

Evan Williamson: Great. We'll continue on.

All right. So what will be evaluated in Stage Two? So I mentioned we'll have a checklist review, and that will go today. It'll also be discussed on the Friday call with the other group of steering committee members. And the final recommendations for that will be due Monday morning at 9:00 a.m. You should have all received the survey today. I sent them via SurveyMonkey and that's how we're going to collect the information for the checklist review. So they'll discuss it today.

And then, (we do have) a call today, you can submit your recommendations. Or if you have some time and you want to think back and read through the checklist or the measures submission.

So, the rest of the evaluation throughout the rest of the process will be on the scientific acceptability. So those are the measure specifications and the testing results. And then our feasibility, our usability, and use criteria.

So here's our Stage Two activities and timeline. You see we're having orientation today and Friday. We'll be having the conference calls on April

3rd and April 8th – those should be on your calendar. It'll go out for a public and member comment throughout May. And then we'll have member voting in July.

They're going CSAC review in August; and following CSAC review, we'll go out for board endorsement. And then there's an appeals process, which is the 30-day process where appeals can be submitted for the measures that were endorsed.

OK, so, we'll be evaluating seven measures in Stage Two. These are the measures that were approved during Stage One and that the developers decided to submit for Stage Two. You see we have five maintenance measures. So they're measures that we're previously endorsed and are going through the maintenance process. And two new measures. So measure that were new to stage one and they're now going through the full review for initial endorsement.

A kind of an uneven split between GI and GU. We have six GI measures and only one GU measure. So we're relying on your expertise in that area for your individual measures, whether be GI or GU.

Anne Pelletier-Cameron: This is Anne Pelletier-Cameron, I just have a quick question.

Evan Williamson: Yes.

Anne Pelletier-Cameron: So, out of all the measures, that, at our initial meeting we had reviewed. So you're saying that some of those the developers have decided not to move on, and that's why they're missing from this list?

Evan Williamson: Yes. Well those are the measures that were approved as a concept and they weren't fully specified or tested yet. And between the time from Stage One and Stage Two, the developers have enough time to fully specify and test it, so they weren't brought back yet. They're still an approved concept and can be brought back for the Stage Two review.

Anne Pelletier-Cameron: OK, thank you.

Male: But some were not put forward by the CSAC.

Anne Pelletier-Cameron: OK, perfect.

Evan Williamson: Any other questions on that? All right, great. I'll move on.

These are the list of the measures. I'm sure you're all familiar with them from reviewing them for Stage One. We have measures from ActiveHealth, from the AMA-PCPI, from Quality Quest, (ARC), and we have – these are the GI measures. And then for GU we also have the measure from NCQA.

Johannes Koch: Hello, this is Johannes Koch, joining late. I'm sorry.

Evan Williamson: Great. Do you want to do the disclosures?

Suzanne Theberge: Yes. Thanks for joining us. If we could just ask you real quick to disclose anything that might be relevant to measures that we'll be discussing today, that would be great.

Johannes Koch: Myself, correct? Yes. Johannes Koch. I'm with Virginia Mason and I don't know that it's relevant, but I've always disclosed that I'm part of the AGA Political Action Committee.

Suzanne Theberge: Thank you. Have we had any other committee members dial-in that haven't introduced themselves yet? OK. Great.

Evan Williamson: Excellent. So at this point, I will turn it back over – I'll turn it to Ashlie. And she will continue with the presentation.

Ashlie Wilbon: So, OK. So, the next part of the call, we're going to focus on reviewing the measure checklist. And this is going to be primarily facilitated by Chris. But I did want to give the group a little bit of background on kind of, how we got to this point and what your, kind of, task is for this portion of the call.

Like Evan mentioned, if you are sitting at a computer and have access to the email that was distributed from SurveyMonkey, to open the survey and kind of, you know, complete the questions for each measure as we go. I think you'll find that it will be useful. So when you go off the call, you don't have

to go back and try to remember what was discussed. So, we strongly encourage you to do that.

So, again, this portion of the call is really to determine which measure should be put forth for full committee review. If you recall – just a little bit of, kind of, history on what happened in Stage One. We reviewed the concept for each of the measures. The Steering Committee made several recommendations to developers for how to improve or refine the concept and that essentially documented each of those recommendation for each measure and provided them back to the developers at the end of Stage One. So that they had a record of the recommendations and what they should be doing before the measure came back into Stage Two.

As, kind of, in addition to the checklist, when those measures went through CSAC, there was a very strong – I'm trying to think of the word – I guess, even stronger recommendation to the point of wanting – it required that the Steering Committee recommendations should be met or at least responded to, in some way, from each developer. So that, there was some rationale or justification on why either checklist items – how they were considered, or why or why not they were not met at the time of submission to Stage Two.

So the purpose of this process is really (bring) the committee through the checklist for each measure. To determine whether or not, based on the recommendations that were made, that you feel that what the developer has responded with was sufficient to move forward to Stage Two.

So, as a result of the call today, and you submit your votes, there'll be another orientation call on Friday (inaudible) checklist. Staff will compile those votes and whatever comes out from those surveys will determine which measures we actually discuss on the call on April 3rd and April 8th. So, if there are some measures of the seven measures that the committee vote should not move forward to Stage Two, then only whatever that subset is will be the focus of the remaining calls.

So, again the measures that we'll be reviewing today for the checklist are all approved as concept. So if you decide that they are measures that you don't

feel like should move – you don't believe should move forward to Stage Two and have a full review against the remaining criteria, they will still remain approved concepts. And the developers will still have the opportunity to bring them back into our process when they feel that they are ready.

NQF (inaudible) back, we don't have a GI/GU project scheduled at this time. We are – we'll be working with developers to determine at what point they can bring their fully specified measures back into the process. Into either future projects against other committees. We're having – we're doing a lot of internal work right now and we're working our process. So we're definitely committed to ensuring that developers have an opportunity to bring their measures back.

So, as an aside, if you decide that, you know, the measures aren't ready for Stage Two, it doesn't mean that they die. It just means that they're not ready at this time for review against the remaining criteria.

Male: Can I ask – basically all these developers, to get to this point, have had to work with NQF staff for technical assistance. So, we're looking at today as their work-product after work with NQF, is that correct?

Ashlie Wilbon: That is correct. And yes, that is correct.

Male: So, it's possible that we're getting their, sort of, "best effort" and if we see that it's not sufficient then that there's maybe finding something went wrong, we have the organization approaching their measure, is that right?

Ashlie Wilbon: It's possible, yes, that is quite possible.

Male: OK.

Ashlie Wilbon: Yes. Does anyone else have questions, kind of, about the checklist review or the implications of the review? The purpose of what we're doing here? OK. So.

Jenifer Lightdale: Before you move on, this is Jenifer Lightdale, I decided to join the call a bit late, sorry.

Ashlie Wilbon: Sure.

Johannes Koch: And, I'm sorry, this is Johannes again. Could you just remind me where the SurveyMonkey is? I'm lost in the flood of emails here, sorry.

Ashlie Wilbon: Sure.

Evan Williamson: Yes, it was the survey this morning around 11:00 a.m. It was actually sent from my email address.

Johannes Koch: Ah, very good. OK. Thank you. Sorry.

Suzanne Theberge: Jenifer, thanks for joining us. Before we get started, could you, since we'll be reviewing measures on this call, we need to do a quick disclosure of interest. Can you let the committee and the developers know if there's anything that you need to disclose.

Jenifer Lightdale: I don't have anything to disclose.

Suzanne Theberge: Great. Thank you.

Ashlie Wilbon: OK. So, with that said, for those of you that are sitting at a computer and have access to webinar, there's a slide on the screen now with two questions, which should mirror what you will see in the SurveyMonkey, which will ask you whether or not you believe the developer had adequately responded to the committee's recommendations and concerns that are indicated in the checklist. And if not, we're asking for just a brief explanation of, you know, why you believe it has not been adequately responded to. And then the following questions is whether or not you believe this measure should be considered in Stage Two, and evaluated against the remaining criteria.

So, pretty brief, but we will be looking for some discussion around, you know, why you do or do not believe that should happen, particularly for the developers on the phone, so they can understand, you know, the committees rationale for their recommendation.

So, with that said .

Zahid Butt: Ashlie?

Ashlie Wilbon: Yes?

Zahid Butt: This is Zahid. I have a quick question.

Ashlie Wilbon: Sure.

Zahid Butt: Can we ask a question outside the checklist or it purely just the checklist and nothing else?

Ashlie Wilbon: We like to focus on the checklist, however, if there are certain questions that you need to ask to give you more context for certain items on the checklist – we do realize that some of the things that we indicated on the checklist kind of require you to go look at the measure. And it may spark other questions that may – you may need to make your decision. So, to the best of our ability, we'd like to keep it scoped down to what, you know, things that are related to what's listen on the checklist. But, yes, just because of the time that we have, we want to try to make sure we keep it focused and concise.

Zahid Butt: OK.

Ashlie Wilbon: Any other question? OK. With that said, I think I'll hand it over to Chris. Maybe with one last note, for those of you that were able to review the memo, we did assign Lead Reviewers. And to the extent that Chris may call on you if you're able to provide some insight on your assigned measure. For those of you who are on the call, there may be some lead discussions for other measures that aren't on the call, but just – for those of you who had been assigned a measure to you on the call and you have some insights to offer, we welcome that as well.

So, I will just hand it over to Chris and take it away.

Christopher Saigal: OK, so, the first one we're going to review is 0098, which is the NCQA measure regarding Urinary Incontinence, assessment characterization and plan of care for incontinent women aged 65 and older. And I just want to first apologize if there's any background noise. I'm in a clinic and it got swamped

with sewage. And since I'm in a public space right now speaking, so I'll try to keep that to a minimum, if I can.

So, I'm not sure if folks can see how the checklist, if we're all on the same page on the NQF website, but there's on page five of this checklist is where 0098 starts. And there are specific recommendations to developer and the developer responses. And perhaps we can review those to start.

So the first one is the committee recommendations with the eMeasures specifications are strongly recommended and the developer comments that – they are essentially saying that they don't have funding to do it. That they were going to do it but they don't have funding. That was the response. And if there's any staff comments about that, you know, please let me know.

The second one is consideration for adaption of patient choice of "no treatment." And the developers said that there is an option if you consider reassessment, and lifestyle modification as "no treatment."

The third comment was expand the age group to include commercial and menopausal population. This is a Medicare age population as measure, and they say, that's a great idea but it's not feasible for them to do that at this time. They may do it later.

Those were the comments I see. Anything else from the staff that I missed in terms of the checklist?

Female: No, that's it.

Christopher Saigal: OK. Does the group have any comments, questions, or concerns about the developers responses?

Anne Pelletier-Cameron: Chris, it's Anne Cameron, I'm sorry but I'm just not finding that the print version of this checklist.

Christopher Saigal: Can someone from the staff direct Anne to the website? I'm on a website basically that – it's the NQF website, there's a link in it.

Anne Pelletier-Cameron: OK.

Christopher Saigal: One of our emails that you received.

Anne Pelletier-Cameron: Please tell me the date that the email was sent, I'll be able to find it.

Evan Williamson: Anne, there are actually multiple ways to get to – I just sent an email out at one o'clock with additional way that doesn't require you to log in. They're posted on our public site, when they went out for comment. You can access those PDF through that. They're under – they're listed under the member comment section of Stage Two, (inaudible). All seven PDFs are there.

Christopher Saigal: And the PDF on the bookmark section says checklist and that can take you to the actual checklist where these comments are. So, does everyone else have access and feel comfortable with what they're reviewing?

Anyone not comfortable with what they're reviewing?

Female: There's so many things on the website, I'm not finding, I'm so sorry.

Christopher Saigal: Understood. If follow the link that was sent out, I guess, by the staff just recently. It's probably the fastest way to do with it.

Ashlie Wilbon: We're going to try to send them out, actually link directly to the PDFs right now, so, hold on for about a few minutes, and you should have those in your inbox. They'll be linked directly to the PDF, so, it should save you a little bit of time.

Female: Thank you.

Christopher Saigal: Actually the NCQA declined to make changes recommended by the committee and it's a combination of funding and resource constraints. And so the decision is whether or not the changes are so critical that the measure should move forward because it impacts the Importance to Measure and Report.

So, I don't have the emails in my Inbox. Anyone else has them yet?

Anne Pelletier-Cameron: Yes. And I went to the one that Evan sent and I'm in the – anyway, under member comment, when I click on the PDF it's just a 16-page or 17-page document and I just don't see where the checklist is located.

Christopher Saigal: So, what I'm looking – so, basically, I think if you click the bookmarks section on the left there, Anne.

Female: The email just arrived from Evan.

Christopher Saigal: OK.

Evan Williamson: And when you open up the PDF, it starts with the memo that includes the checklist responses. So, it might ask you a way up on the first page, but if you scroll down to the second page that's where the table starts for the checklist responses.

Christopher Saigal: OK, that's helpful. So, everyone get that email you'll see then if you click on it, the PDF opens up and then you click on checklist on the left where the bookmarks are and you'll scroll down past 0030, which was withdrawn. And you'll come to – on page five of the PDF the actual comments, back and forth, that we just talked about. So page five of this PDF.

Is anyone not there yet? Everyone is there?

Female: Yes.

Anne Pelletier-Cameron: I have found my way, this is Anne.

Christopher Saigal: So, you know, with that in mind, and there's basically three comments here about eMeasures, option "No Treatment," and expand the age group. They were not able to make these changes. This is a renewal or continuation of an existing measure. Any comments about their responses?

No comments. OK. I will tell you my view, is that it's basically the addition of an option for no choice. I think it's reasonable enough and what they've commented on. The eMeasure lack of specification isn't a fatal flaw on my mind. And expanding age group is important to do, but it’s important to measure in older women as well. So I won’t make it a fatal flaw of this

measure either. So, that's my viewpoint. I don't know if anyone has any other comments about it.

OK. So then, I guess, what you can do is, if your still online, you can vote. Currently, on your SurveyMonkey link, which if you look at that – open that first page with two questions to answer: the first is, if the measure developers addressed the scheduled recommendation test satisfactorily. And if not, please say why. And the second one is, should the measure move forward for a full evaluation in terms of scientific acceptability, usability , feasibility, which we'll all discuss in April. So, those are contingent questions and you either can handle that now or you can come back if when you want to consider it.

OK. So, to move on to the next. The next one on this list, which is opening up now 60622. These are all G.U. measures and these ones are from Active Health. And I can review, on page two – I can review or with the group, the comments and the responses.

So first comments was, was this measure going to include chronic GERD patients. And the developers' response is basically that it is including them in the denominator. Next committee response was exclusion should be clarified in terms of previous malignancy and this is basically upper G.I. study and adults with alarm symptoms which we talked about when we met last time. And the exclusion has been clarified as metastatic malignancy.

The next comment is Barrett's esophagus should be included and they say that this conditions been remove from the denominator exclusions and the (inaudible) will be included if they meet the remaining the criteria, it's not responsive. The measures should be expanded to include patients under 18 as well. And they complied with that. Additional evidence should be provided for evidence criteria and if a – yes, (inaudible) section 1C28, there's no information performance gaps has been provided for us to review. And define testing procedures for the new numerator more clearly. So, that's on the measure testing submission is formal attachment.

Consider also specifying the new numerator at a patient population that would be more broadly impactful. And they've modified the measures so it has two

separate numerators one for the general population, and one for those at high risks. But they haven't tested it and they recommend separating these high risk individuals into noted denominator reporting it separately. I guess they had to back and forth with NQF staff about this.

Staff, is there comments about that?

Reva Winkler: No, but you should see the addendum, I think in red (inaudible) box that last …

Christopher Saigal: Right.

Reva Winkler: Comments?

Christopher Saigal: So basically they're currently casting their – this change. But it doesn't – we don't have any data about how this affects the measure. There's just not enough time for them to do it. And I guess we can the if you want to look at the avenues criteria; this can become this document, right? Under the evidence part there.

Reva Winkler: Yes.

Christopher Saigal: The updated form.

Evan Williamson: The full submission – so the evidence is listed in the docket as well. There should be a bookmark for the evidence.

Christopher Saigal: OK.

So, I've assume that most of people on this call haven't have the time go over back a specific form, in terms of additional evidence, evidence criteria in (inaudible) performance gaps. So, we can either discuss that now or, people can read it over and make their own decisions. Does anyone on the group reviewed that evidence that they submitted? Or have any comments about the performance gap, additional information?

Zahid Butt: So, Chris, this is Zahid, I remember from the original discussion that one of the concerns was that, the numerator statements had gastric emptying study,

as the only test that were qualified as a positive on the numerator. That's not in any of questions that was proposed.

Christopher Saigal: So, what you are saying is that there is an issue that was not raised in the checklist?

Zahid Butt: At least it seems to me, I think there was some discussions that, for instance, if some had a gastric emptying studied, looking for this is looking for GERD patients with alarm symptoms and looking for I suppose, esophageal issues and were actually be counted towards the numerator if they had a single gastric emptying study.

Christopher Saigal: So can we, staff comment on that?

Reva Winkler: I'm not sure exactly what he was referring to but just on the page below that on the PDF, there is if you can read the numerator statement maybe that will help refresh some of our memories. The patients the numerator states that patients who's had at least one esophageal procedure, upper G.I study and that includes upper G.I. radiologic exams, high density barium, with or without the lead film, esophageal or gastric multiple study, gastric emptying study, gastric analysis test, upper G.I endoscopy or upper G.I theory. And it goes on to say, or gas tract to me, evidence of at least gastric or esophageal cancer diagnosis in the last in the past 12 months.

And so, I think maybe if you still have an issue around about it, it might be useful to ask the developers who are on the phone about whether or not that is something that is still warranted. I vaguely remember that in the discussion, but I'm not sure why or wasn't it captured.

Zahid Butt: Right. So, my question only is that, if a patient, who has alarm symptoms, has a gastric emptying study would they be counted as having had that evaluated. That's the implication of this numerator's statement t. And if it is that is true, then I'd like to want to know if it's true.

(Bonnie Veer): Hi. This is (Bonnie) here from ActiveHealth. Can I comment, though?

Christopher Saigal: Please.

(Bonnie Veer): And so, yes. He is correct, it's in the numerator. This was not brought up to us as a concern earlier in any sort of documented correspondence. If it is the request or recommendation of NQF that we removed that, we can certainly do that. That would require, then again, retesting without that piece of the numerator the NQF have that requirement, as well.

Evan Williamson: Well it seems like it's unfair for it wasn't communicated specifically. I don't know whether this is unplanned or not. I'm not an (inaudible) or a urologist. But it seems it's important enough for you to raise it. Zahid, can you maybe explain to the urologist in the audience what the concern is going to be?

Zahid Butt: So I think that this might be – this measure, really is trying to address people who have chronic reflux and have alarm symptoms in the context of chronic reflex, the alarm symptoms really to look for any esophageal lesions perhaps with Barrett's and esophageal cancer and it's kind of that what it comes down to

Christopher Saigal: OK.

Zahid Butt: And my concern is that, they would be numerator statements of flash and if a patient had just gastric emptying study and or no esophageal evaluation at all, that they would be counted as having met these measure. And I think that not be accurate or certainly be interested in thinking and ask for the other G.I. report. And I know at least a couple of the top level G.I on the call.

Christopher Saigal: OK.

Jenifer Lightdale: So this is Jenifer Lightdale G.I., actually I'm a pediatric G.I. Maybe the answer is what we're looking at, data on performance gap. Instead of using the word upper G.I. study, if you could change that to Evaluation of the Esophagus. You know, appropriate evaluation of the esophagus. Zahid, one thing obviously is a little bit confusing, is that there is a lot of different alarm symptoms; on the one hand, most of what is asking for answers if it is cancer. But some could have a picture they are looking for.

Zahid Butt: Correct. Right. I think this is really trying to say, that if someone has GERD and alarm symptoms, they should have their esophagus evaluated.

Jenifer Lightdale: Evaluation of the esophagus. Instead of using the term upper G.I, which is just not quite big enough. Broaden it.

Christopher Saigal: So the thing is, is that we have to, I mean we kind of specify these tasks for their new numerator.

Jenifer Lightdale: Right.

Christopher Saigal: And if we're going to ask them to change that that would require retesting. And so my concern is that we if this is a point that was raised in our initial meeting and wasn't communicated then, then that's a problem in our end. I want to be fair to the developers, maybe the staff should comment about that whether what we should do about that.

Reva Winkler: (Inaudible) endoscopic evaluation all the way through, until that last moment when it says upper G.I.

Evan Williamson: Right. So, basically upper G.I. is not is an inadequate, if it went down alone, inadequate evaluation for alarms.

Zahid Butt: The gastric mortality study by itself as a single test would be inadequate evaluation of alarm symptoms in GERD.

Christopher Saigal: OK. In terms of the developer that's on the call, I guess changing this would require retesting?

(Bonnie Veer): If that's, from my understanding, the NQF does require us to retest and revalidate and if we're making a drastic change to our measure. And again, because this is so late in the game, we would asked the NQF to come up with a consensus agreement on what exactly they we would like to see in the numerator based on expert opinion across the committee. And give us adequate time to put those changes into place after we've discuss them internally and adequate time to test, I think right now it's really late.

Christopher Saigal: Of course.

Johannes Koch: Can I ask you a quick question? This is Johannes Koch. In this evaluation, is there a way to identify or a inexistent data? How many patients have had only an esophageal or gastric mortality study and only gastric emptying study? My suspicion is that these are commonly done in concert with an endoscopy which we're actually wanting to measure. So I'm wondering by the recreating a bar that doesn't need to be done or rather than having it retesting the entire thing without that?

(Bonnie Veer): For us to do that, it would require a very big effort. Exact level analysis. Which you can only fire off one data point, as I supposed to two, if that's what the request is. But, certainly that is something we can look into. But then again it would take time and resource unfortunately.

Christopher Saigal: Sure.

Female: Can I ask a question? Sorry. I'm just looking at 1B.2, 1B.3, right. So you have 392 patients in your numerator, and of these 260 had upper endoscopy and that's written in 1B.2. And then I think where Zahid is bringing up is in 1B.3, the 260 had an upper G.I. If you just change that wording to upper endoscopy (inaudible) the same (250) patients.

(Bonnie Veer): So we're referring to two separate parts here. Because earlier, I thought we were referring to the numerator statements. And now we're talking about these data and the performance gap and 1B. So if that's what you're referring to, we can certainly change that statement easily enough.

Female: I think that's what clarifies it, what you're talking about is what we think you were talking about.

Christopher Saigal: I think it should be more, if we can conceptually, it's inadequate to have a just a upper G.I. or motility study to rule out warning symptom from cancer and the numerator. It's possible given the data that everyone did have endoscopy, given that there using the same terms, but going forward, the median populations doesn't have endoscopy and motility study.

Female: OK.

Christopher Saigal: So, you know, (inaudible) staff, in terms of the way the panel is, this sounds like this is a concern that Zahid that he brought up that was not communicated to the developer and I'm not sure how to manage that in terms of the timeline.

Taroon Amin: Chris, this is Taroon, I have a suggestion for you and the committee. The first is that, as we describe in the two-stage pilot, the developer could just have a level or stage 1 approved concepts and they would have another opportunity in the future to submit the second stage of testing. So it would not be necessary at this stage two of the evaluation, but we would offer an other opportunity in the future for them to bring back, should the committee recommend some changes at a future date.

The second thing I wanted is just to make sure that we keep in mind, is that, the evaluation and specification really is a stage two issue. We are looking at precision and ability of the specifications of how the measures area actually designed. Some the issues that being raised are really issues in choosing the specification which are certainly under the evaluation of stage two.

So, you know, we want to keep things running clean in terms of evaluation of the concept. And specifications will be evaluated fully in stage two. So it's not that there we're reevaluating the measure in stage two, but we're really looking at how the measures and designs, and making sure that the way the measure designs is valid and reliable. And it seems that maybe some of the issues that are being raised are related to the specifications. However, I will deal that back to you guys, in terms of how you want to handle that.

Female: Just to piggyback on Taroon, I mean think you know, Bonnie mentioned in there, checklists if they are still testing parts of the measures. So it is still somewhat correct and again as I've mentioned, we'll be getting another checklist review we will make opportunities for the developers to bring back the measure within the eighteen month period based on a being approved concept. So that is still on the table of measure does not if you guys don't feel like it's ready, there are other opportunities for them to, you know, continue to pursue endorsement.

Zahid Butt: So, I think I'm OK with that in that context.

Christopher Saigal: OK. So it sounds like the other comments to consider are whether the evidence that was additionally supplied is adequate. And evidence on performance gaps is adequate. So, I would suggests, that we no one's really read that. We did that individually and make our decisions and on my view in terms of everything else, we've got a nice response developer in this last matter, is of importance or maybe more into the scientific evidence or specifications around usability or something like that we can use in our next call.

Any other comments?

OK. So, then since we have an hour left, let's go to the next one which is 0635. So please open that one up. And to go to the checklist there, which is on page the first page which looks like so, this is Evan can see that list here of questions, let's see this checklist is a little different in terms of how it is look at – anyone from the staff, can you comment on the way that this is – all these blank staff comments?

Female: Yes. That actually was blank. They either – this was the wrong document I apologize for that. It should actually be the checklist that's in there.

Christopher Saigal: OK. I see a bunch of blank comments.

Female: Yes. Let's skip over that and go to 658 and we'll come back to 635.

Christopher Saigal: All right. So, we're going to 658. Open that up please and this 658 is surveillance on the polyp surveillance and make sure appropriate follow up interval for normal colonoscopy and average risk patients, it's an overuse measure. So there were several comments here. Is there – I don't if there are anyone on the call, was the primary reviewer for this one to be getting overall comments?

Johannes Koch: It's Johannes, I'm just seeing this. So, I'm scanning it quickly.

Christopher Saigal: Right.

Johannes Koch: I mean I think the criteria for waiting 10 years on an average risk is a pretty well defined criteria. So, I think there's – the whole thing is a little less controversial than what we get into polyps and polyps surveillance which is the next one.

Christopher Saigal: OK. So, there's some specific comments here. Essentially one of the, I mean, we could go into them individually for now. I haven't read them at length myself but basically the first one is rather than measuring with (inaudible) recommended, considers (inaudible) the measured look at colonoscopy performed and then do a (inaudible) where the last colonoscopy was performed in the last 10 years. So basically it is changing how the measures is specific to look at actual procedures used versus the recommendations. And that brings you closer to impactful outcome measure.

And their response, is that they the client can do that because of it turns of burden on referring physicians and who basically had – not the patient waiting for 10 years. That's my reason for that one.

And then the other one, the other comments here are that patients aged 50 years or older, receiving a screening colonoscopy at least or mentioned repeat colonoscopy in one year due to the bowel cleansing. And they said this is an exemption built in for medical reasons for not recommending it, the contender (inaudible) so they are going to report the exemption rates basically. So that's how they are handling that situation.

Consider adjusting the upper age limit for older patients including inflammatory bowel disease and better to find above average risks. So they refer to guidelines from the ACS regarding what high-risk means. And they found out that is in this document those with the personal family history of colorectal cancer, adenomas, inflammatory bowel disease or (inaudible) syndromes. And they have to – with a recommendation for upper age limit. And they will bring these cutbacks to the extra workgroup and measure due for review and enhancement.

So, I guess that means that they're not going to do it now. So does that mean anything from the staff?

Reva Winkler: Yes.

Christopher Saigal: They're going to hold off. When a measure is due for review and enhancement that means within the organization that I look at it, if they are trying to look at it again I guess but they're not going to address it right now. And then (inaudible) specifications whether the access are included in the denominator or should be calculated as a separate measure. (Inaudible) not included in the denominator and they're going to report them side by side.

OK. The developers should expand on the available evidence and details of meta-analysis that demonstrate body of evidence (inaudible) and if they did their best and it's been approved before, basically. And eMeasures specifications are present within stage two. So essentially, I guess, the biggest change recommended was involving an actual colonoscopy used as a post recommendations. And there is too much burden to do that. That is more on the feasibility question, I think, than anything else, so we can get back to it, in that time point stage two in my point of view. And any other comments about those responses and the rest of them?

OK. So, in that sense basically people can vote in the SurveyMonkey link or they can defer and think about it some more.

All right. So let's go to 0659. Let me open that up.

Female: 659 is actually in the same checklist as 658 because we sent one checklist per developer. So, if we just scroll down to the bottom of the checklist that you were looking at, you should actually see 659 as well.

Christopher Saigal: OK. So this is actually the inappropriate use one for colonoscopy intervals for patient with (inaudible) polyps. All right. Any comments? Just the group (inaudible) on this.

OK. So, the first comment, as a the developer should expand on the available evidence on the details of the meta-analysis (inaudible) similar to last time and they did their best to do it in (inaudible) first evidence basis. Submitted it. 659 (inaudible) the measure does not match your recommendations in the evidence three plus years versus five years and then consider how those to be

aligned to measures – essentially the measures that have been in (spaced). They argued that they are consistent with evidence based guidelines (inaudible). I don't know if that's true or not. Can anyone comment on that?

Johannes Koch: Well, this is Johannes again. The challenge comes, this is not just whether you have an adenoma, the size of the adenoma and what your previous findings, I think the challenge in it is that there are has guidelines, the problems is in aggregating all that information in one measure, right. So you have this colonoscopy but actually don't know what their previous colonoscopy report might have been and you don't know what other risk factors they may have.

So, I recognize that this is a challenging area now. If you have an adenoma, it is going to be somewhere in three to five. So, in some ways, when there's really some big fish to fry, you can worry about whether it's four years or three years or five years. I'm not certain, but their reflecting on the complexity of doing the space time guidelines when some of the information may have required, you know, what the previous endoscopy was. And some of that can be very difficult to define.

Christopher Saigal: So, essentially it can't get that specific and with the available data and they're going to lump everything together basically.

Johannes Koch: I think they're going to take, yes. I think they're going to take the three to five years and treat it as essentially as one group. So it eliminates people who has cancer or high risk polyp. And it eliminates the other end of the spectrum which is the previous metric, which is people with the normal colonoscopy. So this is the intermediate risk and you'd want to make sure that people aren't recommending 10 years for people who are at intermediate risk. And you want to make sure that they're not recommending one year for people and in whom that it would be inappropriate.

Christopher Saigal: OK. Well, as a urologist, this is definitely something I'm going to defer to my colleagues in G.I. in terms of, you know, how that would play out. And then in terms of how it affects the importance of these measure. Any general comments about that?

Zahid Butt: So, This is Zahid. I think I kind of agree with what your Johannes sort of summarized based on their responses.

Christopher Saigal: OK, OK, well then, with that on mind, the group can either go to the SurveyMonkey and vote unless there are comments. And you can go to the next one 2056.

All right. This is the colonoscopy quality index. Remember this is a very interesting debate that we went through. When I opened this link I'm just getting the (inaudible) bookmark, all right.

Did you guys send around the checklist for this?

Male: It doesn't seem to be.

Christopher Saigal: I don't see a checklist on this.

Female: This one is at the end.

Evan Williamson: It was included in the appendix. (Inaudible) separate it out, so it is included at the end.

Christopher Saigal: OK. All right.

Female: (Inaudible) combined PDF.

Christopher Saigal: OK. So, we actually had a subsequent call with a staff with these folks to address some of their concerns. And there has been going back and forth. I don't see someone tell us where the checklist is on these documents specifically what page?

Female: (Inaudible) if you look at page 98 of the PDF, 98 until 103, I think it's where everybody wants to look.

Christopher Saigal: OK. All right. So all right, so I see page 98. There's one

Female: Ninety-eight until 103. You'll see it in those pages.

Christopher Saigal: OK. I see they're just turned sideways, let me rotate these. OK. So, basically the committee on item two, said that this component shall not be completely composite. It's basically a standard medical practice whether (inaudible) cancer or polyp surveillance. And this is basically given the first credit for a standardize medical risk assessment prior to colonoscopy, is that correct?

OK, so basically they said that the colonoscopy is a (basic) procedure, and assessing ASA is important and should be done. So, essentially, you know, it's a question for the group about whether the assessment of anesthesia risk (private) colonoscopy warrant its own measure, or whether it is standard of care, that should be implied.

Any comments about that?

OK. So basically they didn't feel the change is necessary. Standardized assessment of bowel prep and basically we said the document in the bowel prep isn't really an indicator of quality and there should be some other way to tell whether it's high quality for example, whether the procedure was scheduled for poor bowel prep. And they basically also said that it's important quality step and they feel that it should be included within the degree with us. So, any comments on that?

Male: So Chris, just quick clarification here. The task, just so that we are clear on the task, is to you guys already had a lot of discussion on these elements, it's to open up the discussion on whether or not the recommendations from the committee were addressed by the developer – yes or no. If not where they're essentially, if they're big enough issues to stop it from going forward to the stage two. And so, just to quote it, for discussion, that's the goal. It's the not – I mean the goal would not be to try to rehash of each of these element again because you guys have quite have spent quite a bit enough time on that.

Christopher Saigal: All right. So, that's fine. I just want to show the committee sort of understood what they said and what the responses were (inaudible). Seeing it for the first time.

Male: That's great.

Christopher Saigal: The other two things we talked about were regarding (inaudible) whether or not adenoma detected at that time and they recorded essential pop information. And again, this is – our G.I. folks felt strongly about and basically they don't feel – they don't agree with that.

And then withdrawal time was another thing we talked at length, whether is it important or not and they felt that essentially it is important and they have evidence that they cite about that.

There are complications item eight, they did agree with it including the timeframe and component title (inaudible) of this component. And item nine, is a perfect follow up recommendations and the timeframe for this and they don't collect that information and – I guess they didn't really feel that was important to include.

So in general, they didn't really find the committees comments were worth incorporating. Again, I would (inaudible) comments from G.I. colleagues about, you know, I think that we really did have a wider discussion, thought that they were important is my – for the fact is. Any comment about their responses or?

OK.

Male: Well I mean I would have hope for a little bit more given the length of the discussion of what I thought was some of the quality points that were brought up not from just us, but also from the three societies that wrote a letter. So I don't really know what to say when they say, we listen to you, but they are not going to anything.

Christopher Saigal: Yes. Yes. I hear you.

Bonnie Paris: This is Bonnie Paris from Quality Quest. Can I make a few comments, please?

Christopher Saigal: Sure.

Bonnie Paris: OK. Well first there were several comments that were about elements that are included in the measures which was something we can use for several years. You know, that they are standard of care issues such as having a standardize assessment of bowel prep and medical risk assessment, et cetera. And those are the things that, you know, we are looking at the measure overall and looking at all the historical data that we have to see. You know, if for example, each and every single element is 100 percent correct all of the time, and that really doesn't contribute to the overall outcome of the measures.

So, you know, for example, if the standardized medical risk assessment is performed and documented appropriately 100 percent of the time. You know, that's something that could be removed from the measure potentially without affecting the outcome. But right now we don't have the data on that and so to pull out elements with the assumption that yes, this should be done all the time. But it would create a fundamental change to the measures.

So for those, we do take your advice seriously and we are doing statistical analysis until we get the feasibility of pulling out those different components. And so we can see the results of what that would look like. We can't just change the measure and then all the things that page two submission is based off of the experience that we've had with the measure as it has been stated.

With item six, the central polyp information being recorded; this particular measure is looking at the colonoscopy procedure, and at the time of the procedure we look out whether or not certain information is collected about the polyp and for a separate measure we do look at adenoma detection rates. And for that measure, we use whether or not there was an adenoma detected, but at the time of the colonoscopy exam, we don't have the pathologist's report that confirms whether or not it is actually an adenoma so that's why that couldn't be changed.

So you know I think in some of them it's kind of a disagreement in philosophy but also, you know, if we were to remove several of the components within our measure then it would be a fundamental change to the measure. So that's a drastic change to the concept, and we don't you know have that experience just pulling the things out, changes that can (tap) the measure.

And then also with item nine, the appropriate follow-up recommendation, we don't look at the timeframe, you know, for from the time of the exam, so the patient is given the recommendation. So we're not collecting that, you know, like it would take two days or a week for the physician or physician's office to get back to the patient. We're not collecting information on that and tying it back but we do is that, what was their recommendation, what they're told to follow-up in 10 years, or what they're told to follow-up in two years; and whether or not that was appropriated in the findings of the exam.

Christopher Saigal: OK, thank you. Go ahead.

Bonnie Paris: I just also wanted to say that overall, I mean, we really do appreciate the inputs that you have given us on the measures. So, thank you.

Christopher Saigal: Sure. So it sounds to me like the developer is concerned about changing the measure regarding some of our recommendations without understanding whether the components are tapped out and then used in their community – they want to keep using it. I suggest that, probably, if you have data that would be informative for us to try to see to go to see that at some point, but right now we can only look at how responsive your comments are to our concerns. Yes?

Philip Schoenfeld: This is Phil Schoenfeld, I apologize, I was going to be on the Friday call, but I got a chance to get a way to join in part. Hopefully, I haven't missed too much on all the colorectal cancer screening stuff.

Christopher Saigal: Oh, hi. Actually, we're actually we're …

Male: You came at just the right time.

Christopher Saigal: Yes, but we don't have to rehash all the issues. Basically, there is just – the concerns we have at this point, the developers aren't ready to really respond to them for a variety of reasons. So, they're going to work on that in terms of some other data development on their own. And if there's some other comments about this, we can probably move on. Any other comments about this?

Male: Chris?

Christopher Saigal: Yes.

Zahid Butt: This is Zahid, I have one question from the developer. In item number eight, the title was changed to more accurately reflect that it was the procedure – intra-procedure, or you know, the complication were at the time of the procedure.

In item nine, where it says, "appropriate follow-up recommendation" and in the previous conversation, we sort of did see that the timeframe determines the appropriateness. I'm just curious why they were not able to change the title to just "follow-up recommendation given." Because that would be the more accurate thing in this particular component. Because the appropriateness of that follow (inaudible) …

Christopher Saigal: Right. Well, that's question on the developer, I don't know. But I missed that – it's an open question. I don't know how to answer that.

Taroon Amin: Just so I distill – just to clarify and again, I definitely don't want you to rehash it, but it is the decision that – because I can have comments if needed; but is it the decision on the developer's part that they're not ready to respond to our comments again, so we're not going to move on to stage two at this time?

Christopher Saigal: No, they have responded and they have a rationale for everything that they say. And the rationale is listed and I don't know if you'll get a chance to look at your …

Taroon Amin: Oh no, I've read it.

Christopher Saigal: OK, you have. So basically with their comments our decision is to note whether they're responsive to our comments and whether their comments are valid in terms of their response.

Taroon Amin: So then, I just have one question. Quick question, then, which is should we prepare to discuss as a committee now whether or not they were appropriately

responsive before we vote, or that's something – I mean it's – we'll double check on that.

Female: Yes.

Christopher Saigal: That is now. If you make comments to make that, then we're all ears.

Taroon Amin: OK. I mean my feeling, quite frankly, is that actually they, A, did not respond appropriately to our comments, that with the exception of changing the complication rate or their definition of complication rate that basically all the feedback that we gave them did not lead to any significant or important changes in their indicator and I'll be blunt in saying that I was actually rather frustrated based on the time we as a committee put in to give them feedback about the difference between process indicators and quality indicators, that they basically ignored virtually every specific feedback we gave about the fact that the key quality indicators deal with the indication for colonoscopy being properly documented, that adenoma detection rate has to be properly documented and that the appropriate indication for follow-up has to be properly documented.

But that thing such as specifying in ASA class or simply specifying whether or not a bowel prep quality – whether or not a bowel prep quality was reported; if reported what is the actual quality of the bowel prep? You know, my opinion being strongly being the one who did most of the work on this that they did not respond in a substantive way to our comment, except to say they felt that they didn't need to make changes to the instrument in response to our comments.

Christopher Saigal: I agree with that. Right, so that is probably a fair assessment in my view too. The developer have some plans to look at their data again to see if there's merit in their mind as to whether these are important comments. I'm sure that will be interesting for us to see in some point in the future.

Taroon Amin: OK. I don't want to belabor the case since like I said I was planning on being on the call on Friday. But the developer is not proactively stating we're not prepared to change. The developer is saying that we're not planning on

making any changes to our instruments based on our application. That's the way it looks like to me, and that we wanted to move on to stage two.

Christopher Saigal: Correct. That's where we are right now.

Taroon Amin: OK. Well, I guess I've made my feelings there pretty well known about the fact that I understand the developer feels that responded to our comments that much I appreciate, just didn't feel that it's necessary to make any changes to the instrument despite our recommendations.

Christopher Saigal: OK, that's about right. OK, so with that people can either use their SurveyMonkey vote or they can think about it and vote later.

And then we're going to the last one here which is the G.I. mortality rate from a hemorrhage; 2065, that's an (ARC) measure and checklist is – let me get the checklist up here, the issue I think here was one of sensitivity versus specificity of the numerous definitions. And the first issue was whether or not a primary diagnosis of a upper G.I. bleed was a good idea, whether they should include other secondary diagnoses, and the developer basically responded saying that when you do that, you begin to lose specificity of what caused the death or re-admission. I think on – (turn) on page three, it starts, so that was their concern, essentially. They are going for tightly specific measure where you can really ascribe input into the upper G.I. bleed or lower/upper G.I. bleed and then they feel that's a benefit compared to make it more sensitive or less specific. Any comments about that?

OK. And another comment that we had was concern – stratifying by upper versus lower G.I. bleed or esophageal versus lower G.I. bleeds and they had that you can do that by using a denominator stratification, so that is I guess is possible. So, if there's any overall comment about this, they took our comment seriously. It's I guess our question of opinion as to whether it is better more to be more sensitive and more specific. Any other G.I. comments from our G.I. colleagues on this?

OK, so then we can either vote or consider it and vote later. There's actually one more, isn't there?

Evan Williamson: Yes.

Christopher Saigal: And where can we find that?

Evan Williamson: There's another email, 635.

Christopher Saigal: OK, another email for the last measure that came through. OK. So first 635, opening this up. OK, someone's talking in this – (inaudible) this one, I don't want to be – have so much background noise.

Suzanne Theberge: All right. We're just going to put it up on our screen share here. Just give us a second.

Christopher Saigal: OK, I can start that I think (inaudible) …

Male: What I got is 622, Chris.

Christopher Saigal: 622, that's what we're looking at?

Reva Winkler: It's at the bottom of the page; 635 is at the bottom. Again, it was the same developer, so we only sent them one checklist.

Christopher Saigal: So what we're looking is page – it's on page three, is that right?

Male: OK.

Evan Williamson: Correct.

Male: OK, I see it.

Christopher Saigal: So we're looking 635, (inaudible) hepatitis A vaccination. It's only three of them, if it is difficult to hear me, I'll discontinue.

So our first comment was the numerator was inconsistent with the title of the measure. And containing the title of the measure too, more closely aligned to the measure of focus, and they say if the committee is referring to the inclusion of antibody testing on a numerator, we plan to remove this test, a numerator is no longer allowed to test to be sufficient to complete the numerator. And is that what we meant?

Male: Yes.

Christopher Saigal: OK. All right.

Male: So, Chris, I have a question about this. They have removed (it) from the numerator and there is an exclusion of patients with prior history of viral hepatitis. And my question is, does that also include a positive anti-body test or is it just a clinical history diagnosis?

Christopher Saigal: Well, according to them, they are removing the anti-body test to be efficient.

Male: No, no, no. They're removing this from the numerator which is fine. My only question is, that in the denominator exclusion, there is an exclusion called "patient with prior history of viral hepatitis," but I just want to clarify the question, does that include the anti-body test? Because …

(Bonnie Veer): This is (Bonnie) from ActiveHealth. No, the exclusion does not include the anti-body testing. Anti-body testing only comes across as positive or negative and not as a value, so we are unable to process that, which is the reason why it was the way it was prior to the NQF request. It was in the numerator, we had to infer that if somebody tested for the anti-body that they intended to do something with that information. But no positive or negative testing results; we can't process that information, unless there is a value tied to it. So, no, it was taken out of the numerator and is not included as an exclusion.

Male: OK. Thank you.

Christopher Saigal: OK. So, then the same question comes up about if the person is tested or positively received a vaccination, then they refer to the same solution? Is that OK, folks? Any comments for that?

Male: So I assume that much of the testing that still needs to be done with the numerator exclusion will be ready by stage two?

(Bonnie Veer): Yes.

Male: OK.

Christopher Saigal: OK. And they are trying to harmonize this with 0399 and so I guess that's what they are – is their response, but they haven't been able to report back on any progress that. Any general comments about this?

OK. So then we can either vote on SurveyMonkey or we can consider for later. I think it's all the measures we had to review. And NQF staff, you want to take it from here?

Suzanne Theberge: Yes. Thank you very much, Chris. We are kind of jumping ahead in the interest of time. We're going to skip over the principles of measurement section which you are welcome to review on your own time if you get a chance, and go right into the measure evaluation overview, which Taroon will be (inaudible). So, Taroon?

Taroon Amin: OK, great. So I believe it's slide 31. So as we – page two, just to give everybody – again re-orient now that we've gone to the checklist discussion. We have the stage one review that focus on importance to measure. There were a number of checklist items for each of the developers. And now we're going to move in to page two.

So the measure will be evaluated against the remaining criteria. So as we look at the four major criteria, there is a hierarchy starting with the importance to measure which you already did in stage one. Stage two will evaluate the remaining three criteria which includes scientific acceptability which is a (model) of past criteria, usability and feasibility.

And I'll go through each of these individually. To evaluate whether any one of the four major criteria have been met, you'll evaluate several sub-criteria and expert judgment here is very important. Each of the sub-criteria helps to demonstrate whether the major criteria was met. So, specifically, how do you know that a measure is scientifically acceptable? Those will be where we'll evaluate the sub-criteria to help assess this. The criteria parallel best practices for measure development; for example, we begin with something as important measure and then later what is feasible.

So you already have reviewed what's important to measure, now we'll evaluate and make sure that they're scientifically acceptable and then they're usable and feasible. It's important to keep in mind that the criteria and sub-criteria are a matter of degree rather than all or nothing judgment, similar to what you did in the first stage in this effort.

So I'll get into the meat here, criteria number two is really where you'll spend the majority of your time in the evaluation in stage two. Here, we are looking to evaluate the extent in which the measure as specified produces consistent and credible results for the quality of care method. So, specifically, what we're looking at is the reliability which is a (month) pass criteria, and this is similar to the discussion we are having earlier today to ensure that there is precise specifications including the exclusions and that there is reliability testing either at the data element level, measure score level or both levels. Secondly, you'll be evaluating the validity of the measure's specifications, so specifically whether the specifications are consistent with the evidence.

Just want to take a moment there and highlight the fact of how this is highly different from what we're looking at in stage one. As you'll know, when we look at the important-to-measure criteria, we are looking whether the measures were conceptually important to measure. Here you're going to actually look to see and ensure that the measure's specifications are consistent with the evidence, and make sure that the way that measure is designed is appropriate and valid.

Secondly, we'll be looking at validity testing at the data element level or at the measure score level ensuring that there's justification for any exclusions and how it relates to the evidence, the risk adjustment approach for outcome measures, which is really limited to just the (ARC) measure, if I remember correctly, ensuring that the measure is able to differentiate differences in performance and whether there is comparability across different data sources and methods; and whether there's stratification for disparities if appropriate.

So again with reliability and validity, will just stress that they are not "all-or-nothing" properties but they're a matter of degree and they vary with different

conditions of using the measure. But it's important to focus on the fact that in order to be valid it must be reliable, but reliability is not guaranteed validity.

So, NQF requires that measure testing – that there is empirical analysis to demonstrate reliability and validity of the measure as specified. This analysis should include issues that's pose threats to the validity of the conclusions about the quality of care, specifically, exclusions, risk adjustment approaches and methods identifying differences and performance.

So, key points to keep in mind as you start to do this evaluation is that we require empirical evidence of measure testing. So this is not sufficient for measure developers to simply demonstrate that coding of the measure is reliable and valid. We expect that there is empirical evidence here and also in terms of evidence for testing for the composite which you'll see as a few different elements (or add-ins) for the composite testing form. And that reliability in the measure and the validity is tested for the measure as specified.

If the measure has been updated in terms of stage one and stage two, we expect that the measure as specified and tested and not the measure concepts. That the measure's specifications under reliability and validity of (site) and evidence. NQF allows flexibility in terms of testing options, we're not prescriptive, but those testing results should be within acceptable norms. And insufficient evidence cannot be evaluated or considered for endorsement. To the extent that if there are measures in this project that do not have sufficient testing, they cannot be recommended for endorsement.

So, while NQF requires testing for reliability and validity, we try to implement several methodologies here to reduce the burden of testing for developers. And what we mean by that in particular, is that reliability and validity can be tested either at data element level or at the measure score. And so what we mean by data element here, just to give you some examples for numerators, we are looking at A1C value to make sure that those data elements are valid and reliable, and the denominator can reliably get this information and identify the denominator population, or at the measure's score

level; so looking at the measure's score and ensuring that the measure's score is reliable and valid.

The testing can be done on samples and although we would expect that the scope and extent of the testing is sufficient for the purpose of evaluation. NQF also allows phase validity of the measure score, the systematic indicator of quality if it's systematically assessed and if, you know, just want to highlight that as well.

So reliability of data elements refers to the repeatability and reproducibility of the data elements, the same population in the same time period. So examples here include, inter-rater or intra-rater reliability of the coding or record abstraction as an example. Or internal consistency or reliability of item instrument, although we don't have multi-item instrument, well actually we do have multi-item instrument in this project. So that's also another part that we would look at.

Thanks for that clarification. So reliability of measure score refers to the proportion of variation in the performance scores that's due to systematic differences across the measured entities. So, essentially, what we're looking at here is the precision of the score and examples of this could include, signal-to-noise analysis which gives you a sense of whether the variation is due to the differences in performance of the measured entities versus just random noise in the measure score.

So, I'll just move on past this slide actually, it's pretty straightforward. So in summary here, what we're really looking here is whether the test is appropriate with the appropriate method used, and we'll ask you to consider whether or not the level, the data source, the type of measure, the topic and potential sources of error, conceptual relationships are appropriate in terms of the testing. Secondly, whether the scope of the testing is adequate; if there is a sample, whether looking at the number of entities, number of patients and representativeness of the testing and whether the results are within acceptable norms.

So, again, NQF requires empirical testing of these measures. So if there are any further questions relating to what we would require in terms of testing, please reach out to us and we'll give you some additional information in terms of our testing task force recommendations and what we require for testing. But it is a fairly high bar in terms of empirical data as the measure specified.

So I'll just move on here to just focus on a few other components here which again really is only focused on one major measure you'll be looking at. But as far as risk adjustment codes, we expect that to include patient level factors, for example, age or diagnosis, severity that are associated with the outcome of interest, and we do not – don't confound with the quality of care provided. So particularly, we expect that the components that are included in the risk adjustment model are at the present – are present at the start of care, so they're not a result of care and then being used as risk adjusters for the outcome. That would not be a positive way to design the measures.

Secondly, that we generally do not include measures that are associated with disparities. For example, (SCS and rates) and we generally don't include structures and characteristics of organizations or clinicians that would be associated with quality. For instance, experience, training or equipment would not be appropriate variables to include in the risk adjustment model.

Oh, sorry, next slide.

Female: Yes.

Taroon Amin: Sorry, OK. So, how do we rate – I just want to spend a few seconds here to talk about how are we rating our reliability and validity. So in order for a measure to be rated high for reliability and validity, it needs to be empirical evidence and testing both at the data element level and at the measure score level. If the measure is not tested, is only tested at one the highest rating it can receive as a moderate and that the major differentiator between high and moderate ratings in reliability and validity. And the differences between low and insufficient in terms of reliability/validity testing is that, insufficient evidence does not have any empirical testing of the measure. And when we look at the overall – so on the next slide, overall, a measure that does not pass

scientific acceptability is either rated low for validity and/or low for reliability – yes, and reliability.

So, for measures rated low in either reliability or validity, it will not pass scientific acceptability and the measure will not be endorsed. It is a must-pass criteria. So I'll turn it over to Suzanne to just quickly go over the remaining two criteria. But I will say that the majority of your evaluation will really be on the scientific acceptability section of the submission form.

Suzanne Theberge: Thanks, Taroon. OK, the third criteria is feasibility – sorry, is there any questions so far before we dive into the next criteria?

Zahid Butt: Yes. This is Zahid. I have just a quick question for Taroon. I think you mentioned that there is guidance on the testing methodologies. Do they also have minimum thresholds in them?

Taroon Amin: We don't have minimum thresholds. We don't set minimum threshold. Again, because many of the thresholds are contact-specific but we generally expect they're within acceptable norms; and those acceptable norms are relatively well-defined in our testing task force report. But they're not meant to be minimum thresholds.

Zahid Butt: OK so, because I was just scanning the 0635 with testing that they are to complete submission that there were several values that were used to justify the results, so I was just, sort of thinking ahead. I wasn't sure whether that was acceptable thresholds; I guess I'll just have to see what guidelines document says whether that addresses the issue or not.

Taroon Amin: I think that there's two different avenues we could use for that to the extent you have specific questions on interpreting any of the results that the measure developers have submitted, please feel free to reach out to us so we can help kind of describe that in a little more detail with our lead methodologist here at NQF during in-person or during our meeting.

The other option is that we really – some of these that are in the panel here there maybe some that feel more strongly about what an acceptable scope or

result for testing. But again, we don't have minimum threshold, I'll just underline that.

Zahid Butt: OK.

Suzanne Theberge: Any other questions? All right. So feasibility is our next criteria and that's the extent to which acquired data is available, retrievable and can be implemented. So, for clinical measures, that would be things that required data element are generated and used or in care delivery – something like, blood pressure, blood test, medication orders, stuff like that. So required data elements are available in electronic health records or other electronic sources. And that demonstration as the data collection strategy can be implemented, so they're not issues of patient confidentiality with timing, cost associated with fees of proprietary measures, and that there's some indication that the implementation has either been tested or are being used.

The next criteria – is there any questions on that before we go into the next one?

Christopher Saigal: Occasionally I had is in the past, we had – this is Chris. I've been in some NQF meetings where the participation and the data – specific registry was considered undue burden. Is that what's still sort of the thought process over here right now?

Reva Winkler: Chris, this is Reva. In terms of you mean requiring to participate in the registry as the only way to collect data on the measure. Is that what you are referring to?

Christopher Saigal: Yes. Feasible only if you're part of the registry.

Reva Winkler: Right.

Christopher Saigal: You're on page – go ahead.

Reva Winkler: Yes. I think that whole area is evolving and registries are much more prevalent and people are involved in them much more. But I still think it's the topic to discuss and to weigh in your evaluation because clearly there may be

barriers to participating in the registry that would preclude some of the people from participating and therefore not being included in measurements. So you do have to weigh that in as a factor, though registries are becoming very much more used and popular.

Christopher Saigal: OK.

Suzanne Theberge: OK, any other questions?

All right, so the last criteria is usability and (you know switches) the extent to which potential audiences are using or can use performance result for both accountability and performance improvement. And in the interest of time, I'm going to just jump past this, this is pretty straightforward as well.

Why don't I talk briefly about related or competing measures because that will be an issue in this project. If you have a measure that is endorsed and then you have another measure that has the same measure focus, the same target population, these measures will be considered competing. And so we do have an issue in this project with the hepatitis A vaccination measure and if that measure, once you reviewed it and if it passes all the criteria and you recommend it for endorsement, you will then need to discuss how that should be harmonized with the other measure.

And as we mentioned earlier, we've ask the developers to talk about that and they will send us some information which we'll share with you before the meeting in April.

Briefly, (we're) use to high, moderate, low and insufficient meeting categories for those as well. And then new versus endorsed measures, we touched on this a bit earlier in the call, but five of the measures that we're looking at in this project has been previously endorsed. So these slides show some of the things that you should be looking for in previously endorsed measures versus new measures. And we do seek implementation comments from our members before we look at endorsed measures where we review and we'll be sharing those with you shortly as well.

Are there any questions? All right. So I just want to go over the next steps. There are a lot of things going on in the next few weeks.

So we have another orientation call on Friday where we're going to go over the same stuff we went over today and do the same checklist discussion. You're welcome to join but you definitely don't have to since you were on today's call. We do need you to vote on your preliminary checklist surveys if you have not done so today. I know quite a number of you did, but if you haven't, please complete those by 9:00 a.m. next Monday morning. We'll let you know on Monday the results of those. We'll combine the votes of both orientation calls and let you know if all seven measures will be going forward for the full review.

So, at this time, you can also start looking at the full measures submission forms. Those are all posted on SharePoint. If you have trouble finding them, let us know. But you should start looking at that and you should complete your preliminary evaluation surveys by April 1st. We sent that link out in a memo that I sent earlier, but we'll send that again next week so that you can have all that information together when we know exactly what measures we'll be reviewing. So you'll need to complete those preliminary evaluation surveys by April 1st and those are all still due at 9:00 a.m. Eastern Time on April 1st.

We'll compile those results and send them back out to you later that day so you have a little bit of time to look over them before the first committee call on April 3rd. So that call on April 3rd, we're going to look at all the measures that are under review in this project. We'll send out an agenda so we know which measures we're going to be looking at. That call is from 3:00 to 5:00 Eastern Time.

We do request that if at all possible that you'll be at a computer during that call, since we're going to be meeting by phone and not in-person, we'll be doing online voting, so we'll need you to be on-line, so you can access the survey tool that we'll be asking you to fill out. Since some of the criteria are must-pass, we need to know right away what your vote is on a criteria so we can move the measures along. And we'll be doing the same thing for the April

8th call and please plan to attend both calls if you can. We need the full committee's input to decide on the measures.

So we have given you quite a lot of information. Are there any questions at this time? So we expect that you may have questions as you start to review things. Once you start looking at the measures, you know, if you run into questions with the criteria, please don't hesitate to call or e-mail any of us, and we can talk to you through that. Here's all of our e-mail addresses and our phone. Again, everything that you need is on SharePoint and I will continue to send e-mail updates as well.

Taroon Amin: Chris, you might want to just open the line up to public comments. If there's anyone who wants to make a – pose comments about today's call.

Christopher Saigal: OK. Yes, so obviously anyone – anyone at this time who wants to make a public comment, please do so.

OK, I don't hear anyone. With that, then I guess we'll conclude it for today?

Suzanne Theberge: Hi, we are unless there is any other outstanding questions.

Christopher Saigal: All right. Thanks everyone for your time and participation.

Reva Winkler: Thank you.

Suzanne Theberge: All right. Thanks everybody.

END