

NATIONAL QUALITY FORUM

Moderator: Alexis Forman-Morgan
November 29, 2011
2:00 pm CT

Operator: Ladies and gentlemen, thank you for standing by. Welcome to the conference. Please note today's presentation is being recorded.

Alexis Forman-Morgan: Welcome everyone and good afternoon to the Surgery Endorsement Maintenance Steering Committee conference call and today the committee will be discussing phase 2 measures as well as discussing related and competing measures regarding AAA measures and reviewing the harmonization plans for some of the prophylactic antibiotic measures and finally, reviewing a newly-submitted measure from ACS measure 1741, consumer assessment of healthcare providers and systems, surgical care survey.

At this time, I'd now like to turn our conference call over to Dr. Morris, our co-chair.

Arden Morris: Welcome, everybody. Thank you for making it to the conference call today. I think our agenda is pretty well lined-up as an overview. We have previously discussed measures 357 and 359, 365 and 366 and we have a few more things to wrap-up on those measures.

We have also previously discussed the AAA measures 357, 359, 736, 1523 and 1534 and a few more things to talk about with regard to those. Also on our agenda we will talk about the harmonization of prophylactic antibiotic measures.

And for the first time we will begin talking about the S-caps measure and then have time for member and public comments and then we'll talk about the next steps of the Lexus. Any questions about that?

Okay, so for starters on the continuation of the phase 2 measures and the follow-up, we were going to discuss these measures in phase 2 and specifically Alexis correct me if I'm wrong but we have some questions for AHRQ don't we as a group?

Alexis Forman-Morgan: Yes. We began looking at these measures on the last call on November 7...

Arden Morris: Right.

Alexis Forman-Morgan: ...looking at measures 357 and 359 and AHRQ noted that they did combine the mortality rate measure and the volume measure together in 359 but still wanted to maintain the volume measure which is 357 as a standalone measure.

And I know that some of the committee members had questions about that and we do have Patrick Romano representing these AHRQ measures on the call and he can give a brief overview of that information and answer any questions from the committee.

And staff did review the measure evaluation form and noted that the combination of the two measures - mortality rate and volume - were only listed in the calculation algorithm section of the measure submission form on Page 9 and so we did have some questions regarding that as well.

Arden Morris: Okay, so specifically Dr. Romano what we'd like to know is what's the rationale for keeping volume as a standalone measure in addition to the fact that mortality and volume are a combined measure.

And secondly and I imagine that that will be the most important part of what we need to know but also want to just stipulate that the combined measure needs to be specified in the description and specification as well as in the algorithm.

Patrick Romano: Okay, well, I'm here and Jeff Geppert is also here representing the AHRQ team so he'll fill-in where I leave off but I can say that the rationale for having the separate volume measure is that they really serve different purposes.

So the mortality measure is a composite measure now of the stratified risk-adjusted mortality where we're now stratifying by open versus ruptured aneurisms and endovascular versus open procedures.

That is a risk-adjusted mortality measure in which hospitals - it's based on a hierarchical regression model - in which hospitals are regressed or smoothed or shrunk if you will back to their volume-stratified means so volume is incorporated only as a prior if you will for the best estimate of performance for small hospitals.

Absent any reliable information about what's the true mortality performance of small hospitals is, we fall back on the mean for all small hospitals of that volume size so the volume is really just in the mortality measure as a basis for the prior distribution of hospital performance.

On the other hand, many users are interested in knowing what the volume of hospitals is because there's interest in directing patients to higher-volume hospitals, selectively contracting with higher-volume hospitals, or having relationships in local communities that reward higher-volume hospitals for maintaining their volume and for having the greater experience that they do.

So the volume measure really has a separate and distinct purpose from the risk-adjusted mortality measure even though the risk-adjusted mortality measure now incorporates volume-specific mortality as part of the calculation algorithm.

Arden Morris: Okay, then I guess that I would paraphrase this to an extent by saying that the data seem to indicate that volume is a better predictor of mortality than mortality is, so I'm incorporating volume into the mortality measure makes a lot of sense. If we could only measure...

Vivienne Halpern: But that's actually not true for endovascular to a certain extent...

Arden Morris: Okay.

Vivienne Halpern: ...because above 10 in the literature you have equal results.

Arden Morris: Is that Carol?

Vivienne Halpern: No, this is Dr. Halpern. This is Vivienne Halpern.

Arden Morris: Oh hi, Vivian.

Vivienne Halpern: Yes, so above 10 has equal - in the most recent literature for endovascular - volume doesn't make a difference.

Arden Morris: Okay.

Vivienne Halpern: It does make a difference for open repairs but not so much for endovascular.

Arden Morris: Okay and does that give you sort of does that give you pause for incorporating volume into the mortality measure for this?

Vivienne Halpern: Well, I'm still trying to figure out that I have to say statistics is not my strong point so I'll admit that up-front but to explain this shrinkage again, how that works for the shrinkage and do you think a volume averaged for all small hospitals and then how would that work on the provider level which may be even smaller?

Arden Morris: I don't think we're going to measure the provider level; is that correct? I thought this was only at the hospital level.

Patrick Romano: That's correct.

Vivienne Halpern: Okay. Dr. Romano, can you clarify that shrinkage process?

Patrick Romano: Well, maybe I'll defer to Jeff Geppert (for that).

Jeffrey Geppert: Sure, so the shrunken rate or what we call the smooth rate is actually just a weighted average of two numbers. It's a weighted average of the hospital's risk-adjusted mortality rate and the volume-specific mortality rate so it's a weighted average of those two things and the weight is a reliability weight, basically a signal-to-noise ratio.

So for smaller hospitals, that ratio tends towards zero and for larger hospitals it tends towards 1 so for larger hospitals there's more weight given to the hospital's risk-adjusted mortality rate and for a smaller hospital, there's relatively more weight given to the volume-specific mortality rate.

Arden Morris: So Vivienne from year to year, I think that just looking at the mortality in a small hospital is not fair in some ways from year to year because, you know, if they do one operation the year before, it's either going to be 100% or 0%.

Vivienne Halpern: Exactly, that's why I have a concern about volumes.

Arden Morris: Right, but if they incorporate volume, then that will make it more truly reflective of the actual mortality if they incorporate volume into the mortality measure. It is sort of fairer to the - it doesn't unfairly - exaggerate the predictable mortality or underplay it.

Patrick Romano: And just to clarify one point so because we estimate these volume-specific mortalities empirically using the empirical data, it doesn't really matter whether there's a huge difference between low-volume and high-volume hospitals or a small difference in mortality.

Whatever that difference is, hospitals will be shrunk toward their own volume-specific mortality rates so for obviously the extent of that movement is going to be greater for the open procedures because there's more disparity in the volume-specific rates, the volume-specific mean rates. Does that make sense?

Vivienne Halpern: Yes.

Arden Morris: Are you satisfied with that explanation?

Vivienne Halpern: Yes, I just, I don't know, as a person who likes to see things in front of them, the whole concept of sort of using empirical data is concerning but I guess that's how it has to be done because you're doing it off of indirect data as opposed to say like a registry.

Arden Morris: Okay.

Patrick Romano: The same method would be applied to registry data and for example the SVS has used similar methods with its registry.

Arden Morris: Okay, and then Dr. Romano can you talk a little bit more about the rationale for continuing to measure volume? I think that, you know, it may be very straightforward to some and not as straightforward to others so continuing to measure volume as a separate measure.

Patrick Romano: Well, it's just that volume has a different policy implication and so for some users, for example for managed care plans or health insurance companies or Medi-Cal plans, Medicaid plans that are interested in developed contractual relationships with hospitals, they would be interested to know what's the volume of the hospitals that are in their market.

Similarly or for entities that are interested in looking at referral patterns in communities or encouraging referrals, it may be useful to look at volume separately.

Many have advocated, you know, specific volume thresholds especially for open procedures and so having a separate volume indicator allows people to apply a threshold that's evidence-based from the prior literature.

Arden Morris: Vivienne or anybody else, any questions about that?

Vivienne Halpern: So essentially you're being the repository of the volume count? Is that what this measure becomes because I don't know that it should be quality measure although I do think that people should have access to that data but I was wondering if there's other ways to have access to that data.

Arden Morris: Is that a question for Dr. Romano?

Vivienne Halpern: It sounds like it's a - I'm just trying to understand it in my head - why it's important. It sounds like it's important for essentially people to be able to just look at the volume number.

Patrick Romano: Well, and it also provides context for interpreting the risk-adjusted mortality numbers because if the volume is very small, then that risk-adjusted mortality estimate is basically just the average mortality per hospitals of that size and so there's really very little additional information about the hospital's own performance that's going into that risk-adjusted mortality estimate.

On the other hand, if the volume is high then that risk-adjusted mortality estimate is more heavily weighted by the hospital's own experience so in general in a number of previous cases, NQF has supported the idea of pairing these measures so that users can have some context for interpreting the one knowing the other.

Vivienne Halpern: I would have to say that, I mean, that may work for healthcare companies and insurance companies but do we really think that the average patient would understand that? I don't think so.

Arden Morris: Let me just say that the, you know, I think that we understand that volume is a proxy for something that we can't fully define and it's often used as a measure of quality even though volume itself is not quality. It's whatever happens in the high-volume hospitals that are doing rare operations and we still haven't gotten down to that granular detail.

But in terms of what a patient walking into the clinic might understand, I do, I think, you know, people in general have common sense and a really common question that I'm sure all of us clinicians have faced is how many of these have you done because people do have a sense that the more you do, the better you should be at it.

And I think in that way - in terms of public reporting - in that way that may make sense to healthcare consumers but you're right that probably the majority of the utility would be for probably for other stakeholders besides the consumers.

Vivienne Halpern: Well, I think that one of the things that's important about the NQF measures is that they are used by a variety of parties for a variety of purposes and there's not, you know, every measure doesn't have to meet every purpose and every audience.

Arden Morris: All right. Anybody else from the committee that would like to comment on these measures or the explanation given by AHRQ?

Peter Dillon: This is Peter. I'm just picking-up half the conversation so I'll stay quiet but the statistical analysis that we've engaged in that I did catch the tail end of sounded appropriate to me so I'm comfortable with them to the extent that I've been listening.

Arden Morris: Okay.

Terry Rogers: Arden, this is Terry in Seattle. I've heard most of the conversation also and I hate to ask such a piss-ant question but who actually pays any attention to this; these numbers?

Arden Morris: The hospital alliance.

Vivienne Halpern: To say the risk-adjusted mortality, who looks at these? Is that what you're asking?

Terry Rogers: Right. From a practical standpoint, do prospective patients look at them? Do doctors look at them or is this will be looked at as part of a process perhaps to adjust payment based on mortalities or shouldn't I be concerned about any of that?

Arden Morris: I think that that's a really good question. I could envision at some point as a sort of a quality improvement movement gains ground that there will be some regionalization of care. Carol might even be able to say a little bit more about that from within the, you know, the paired community.

I imagine that at some point that CMS probably will get more involved with saying who can do what although we're nowhere near that right now.

Vivienne Halpern: Although I have to say there was actually a whole thing. Actually it was on like a 60 Minutes or a 20/20, one of those kind of programs about how when patients don't want regionalized care because they don't want to have to have big operations far away from family support.

Terry Rogers: Right.

Arden Morris: I know there's some more policy-level decisions that would be made at different levels than our committee and what our job is is just basically to figure out do these measures fit the criteria that we have outlined and making a recommendation or endorsement based on whether or not they fit these criteria and in the future I'm sure that more will be done with them.

You know, we may see a butterfly effect. We may see that things linger without having any effect for a long period of time. It's a little bit hard to know.

Terry Rogers: Yes, okay, well, I...

Carol Wilhoit: In terms of the uses, I think I can comment on that. You know, as a health plan we do measure a variety of things for a variety of purposes but we've profiled hospitals, we've profiled physician practices.

We have pay-for-performance programs. We have ACO contracts. We have patient-centered medical home contracts and to the extent it's possible in all of those scenarios, we try to use endorsed measures when we can.

Sometimes there aren't standardized measures and then we have to create our own or adapt something but to the extent that we can use something that's standardized, it's much preferable.

And we are constantly, you know, and there's, you know, centers of excellence programs, you know, so that there's a variety of different programs where the data is used and it's to everybody's advantage if there are standardized measures that we can use.

Terry Rogers: Great Carol, thank you; that was a great answer. I appreciate it. I guess where I was coming from and I won't divert this conversation any longer is that at some point, the kind of information we collect with respect to the services that have already been performed has to move kind of backwards in time and be utilized in a fashion that helps at the moment of care or in the decisions made with respect to even offering that kind of care.

And so that's my hope is that in addition to sort of the post-service analysis and payment and chastisement if that's to come, it would be useful I think if we were to gather all this information and to provide it in such a fashion that it could be enveloped into the decision-making process that happens on a patient-to-patient basis so that's my fond hope and with that I'll be quiet.

Arden Morris: All right. Is there anybody else on the steering committee that would like to talk about this or make comments or ask questions? I think that we need to decide then if these measures are essentially to be recommended. Is that right, Alexis, or do we need to move on to the pancreas measures first?

Alexis Forman-Morgan: We can get an overall consensus but the committee will vote via SurveyMonkey as well after the call.

Arden Morris: On 357-359?

Alexis Forman-Morgan: Yes.

Arden Morris: Okay.

Alexis Forman-Morgan: But if you would like to get a consensus from the group now, that's fine.

Arden Morris: I think we can just move on then if we're going to be doing the exact thing but through SurveyMonkey after.

Alexis Forman-Morgan: Okay.

Arden Morris: So the next two are the pancreas measures and we need to look at the testing that was performed by AHRQ and to give AHRQ a chance to introduce the changes that you have made. Will that be Patrick Romano again?

Patrick Romano: I think Jeff will take the lead here.

Arden Morris: Okay.

Jeffrey Geppert: So the primary change that was implemented at the request of the steering committee last summer was to include the nine cancers in the denominator and to harmonize the specifications for the volume and the mortality measure.

And the request was to present some data that was stratified by both pancreatic cancer cases and non-pancreatic cancer cases so those are the results that were included in the revised submission side.

Arden Morris: Okay, and let's see. Are we going to go through that? Are there questions listed at the top of the measure here that we're going toward?

Richard Dutton: This is Rick. I specifically reviewed these. I'm happy with the data provided and how they have done the stratification.

Arden Morris: Okay. Any other issues, Rick? Anything else that you would like to say about this, any points of conflict that we should discuss or questions?

Richard Dutton: No, I think I'm happy with the measure, Arden. I mean, our biggest objection before was that one of them included all pancreatic surgeries and the other one was just...

Arden Morris: Right, cancer, right.

Richard Dutton: ...something irrational so...

Arden Morris: Okay.

Richard Dutton: I checked it in the data and it looks fine.

Arden Morris: Okay, and then let's see. Alexis, what are we to do now with these measures? Are we to make a decision about them or are we going to be voting on them later?

Alexis Forman-Morgan: Yes. You'll be voting on all of the measures discussed on today's call via SurveyMonkey.

Arden Morris: Okay. That sounds fine. Okay, and now next a discussion of harmonizing prophylactic antibiotic measures.

Alexis Forman-Morgan: Next we're going to review the AAA repeated competing and related measures.

Arden Morris: Okay.

Alexis Forman-Morgan: And so in your package you were given a related and competing measures table of the AAA repair measures and at the very top of that table, we outlined some of the information and listed some things that were different about each of the measures that are under consideration for related and competing as well as anything that was similar.

Arden Morris: I'm really having a hard time seeing the font on this screen.

(Heidi): Would it help - this is (Heidi) - if we perhaps walked through high-level each of the measures?

Arden Morris: It would help me.

(Heidi): If it's hard for you to read, okay. Alexis, do you want to go ahead and do that or...

Alexis Forman-Morgan: Sure. That's fine so looking at the first two that we just reviewed - measure 357 and 359 - and looking at the measure focus, there's a similar measure focus to the current endorsed measure 736, the Leapfrog measure.

However, 357 and 358 if you scroll down to Page - sorry, I'm still scrolling - if you scroll down to Page 28 and you look at the level of analysis, you can see that they're all at the facility level and if you look at the data source of the first three, they all are using administrative claims.

However, if you look at the SVS measures - the two newly-submitted measures to the project - you'll see that they're using registry data in comparison to the AHRQ measures and the current endorsed measure that's using administrative claims.

You can also see looking at the level of analysis for the SVS measures that it's also at the facility level; however, it's also at the clinician level as well.

Patrick Romano: This is Patrick. Could I clarify something?

Alexis Forman-Morgan: Sure.

Arden Morris: Sure.

Patrick Romano: Yes, so just to be clear about the distinction between the Leapfrog predictor 0736 and the AAA repair mortality measure 0359, so 0359 is based on patients level or record level discharge or administrative data and it incorporates a risk adjustment using patient characteristics as specified on Pages 8, 9 and 10 of this table.

The Leapfrog predictor is designed to be collected through a survey of hospitals where they ask hospitals to do their own analysis and to report on their number of cases and their mortality rate using defined specifications but there's no patient-level data that's collected by a central source and therefore there's no risk adjustment model.

So that's the fundamental difference between those approaches. One is designed for someone who has a collection of patient-level data. The other is designed for an organization that's simply surveying hospitals and asking them to report hospital-level summary of their experience.

Arden Morris: I was just reading Leapfrog's reason why they didn't do a risk adjustment too which was basically that they said that there was no difference between doing a risk adjustment and not.

Vivienne Halpern: Yes, they don't have a lot of faith in risk adjustment as a group, so on the plus side for registry data it seems like, you know, data that's collected primarily and not for administrative claims purposes would tend to be more accurate but on the down side, there's no risk adjustments.

Arden Morris: Well, they're not doing even the registry data. They're just doing a hospital - it sounds like they're doing - even less than that by registry. They're just having the hospitals tell them what their survival rate is. They're also looking at survival rather than deaths.

Patrick Romano: Exactly. Leapfrog is simply surveying the hospitals, sending them a questionnaire and then the hospitals submit the answers to the questionnaire and they basically fill-in the number of patients that had the surgery and the number that died.

Arden Morris: Alexis, in this decision-making are we supposed to choose a winner here? 736 is already endorsed, yes?

Alexis Forman-Morgan: Yes. It currently is so first we need to decide is there any harmonization issues or if these measures are competing. If they are competing, we need to select one or indicate why the competing measures should stand and why...

Arden Morris: Why both should stand potentially?

Alexis Forman-Morgan: ...both competing standards, yes, yes.

Arden Morris: Okay.

(Heidi): This is (Heidi). If I could just add a couple more things so 736 is endorsed. The endorsement, the question of whether it should remain endorsed or not is not before this committee.

Arden Morris: Yes.

Alexis Forman-Morgan: The question is on the measures that you have before you so again I think you do need to look at whether or not these are competing measures. I think that in many ways I would say they are.

They have been viewed that way in the past when they've been looked and we have gotten a clear direction from the (C-sac) and the board that in most cases they would like us to determine whether one measure should go forward before another and this is something that the board actually will be discussing even further at their meeting on Thursday.

So the more that you can perhaps come to a decision, now it may be you could say that you believe the AHRQ measures are that better meet the criteria and that would be something that would then move forward when the Leapfrog measures are reviewed and we would have two measures out there.

You may actually also say that you feel that the Leapfrog measure is fine and meets the criteria better than the AHRQ measure. Similar with the SBI and so those are kind of the things that we need you to think through and talk about and then come up with your final recommendations. Hope that's helpful.

Arden Morris: Well that is helpful, thank you. Anybody on the committee want to comment on that or have any thoughts about whether we consider these measures to be competing?

Vivienne Halpern: Well, I think they're definitely competing. They look at basically the same things stated in a different way. The Leapfrog has not, you know, AHRQ has separated out now endovascular SS open and it looks like Leapfrog completely ruptures as any part of it because they specifically say it's for non-ruptured.

Arden Morris: Right.

Vivienne Halpern: But they didn't separate it into vascular versus open so I do have a problem with that aspect of it because that can make a great difference on how you view the outcomes of those patients.

Arden Morris: Is there anybody on the committee that would favor 736 over the others, over 357, 359? Sounds quiet. What about in the other direction? I have to say that I favor 357-359 over 736 and I do see them as competing as well. Anybody have a different opinion about that?

Connie Steed: This is Connie Steed. I concur with you.

Arden Morris: Okay. All right, so although we're not really - it's not our charge - to say too much about 736 at this time, I think that our recommendations will hopefully be brought forth in the future when it's reviewed and if this is the sort of thing that the board is looking for from us, it sounds like basically as a group we're either non-committal or on the side of the fact that these are competing and that 357-359 are stronger than the Leapfrog measure.

Terry Rogers: I agree with that. This is Terry.

Arden Morris: Okay.

(Heidi): One thing, though, Alexis will specifically in the SurveyMonkey ask you to provide a rationale so that we can really write this up so it's very clear to the membership and everyone else and then moving forward as the Leapfrog measure would be up for review in the future, we'll take - carry that forward - as well.

Arden Morris: Okay, and so I would say as the rationale and I'd invite anybody on the steering committee to either add to this or if you have a disagreement to bring it right up.

But I would say in what we've talked about so far as Vivienne said, the AHRQ distinguishes between endovascular and open, that the AHRQ measures risk-adjust and that they have that additional advantage of incorporating volume. I believe that the Leapfrog measure also incorporates volume.

Vivienne Halpern: It does incorporate volume.

Arden Morris: Yes, but does not risk-adjust and also is hospital-reported based on hospital survey and otherwise although if it's labeled as registry-derived, it might seem to carry more weight. It doesn't necessarily provide any additional information over the claims-based method. Anybody disagree?

Terry Rogers: That sounds right.

Arden Morris: Any other advantages to the AHRQ measure over the Leapfrog measure? Vivienne, do you have anything to add to that.

Vivienne Halpern: No, I don't think so.

Arden Morris: Okay.

(Heidi): Thank you.

Arden Morris: Sure.

Vivienne Halpern: Now in terms of the SVS measures, I think they're different because they're specifically looking at smaller aneurisms and open versus endovascular as two separate measures and again the issue is is that there's a lot of if you look at the aneurisms being done out there, a lot of them are done for sizes less than is this threshold for fixing aneurisms or the current guidelines.

Arden Morris: A lot of them are done for aneurisms that are a size less than what?

Vivienne Halpern: Than the recommended - they're done for relatively small - aneurisms.

Arden Morris: Okay, so you're...

Vivienne Halpern: There's a lot done in the community that are done for smaller aneurisms than the recommended size of the threshold and so the SVS concern was that if you're doing these, you better have really good outcomes because you're already, you know, it'll be - the reasons that they're done - at a smaller size are a variety of reasons.

One might be that it's growing faster than the expected rate. One might be the appearance of it and, you know, there's usually some kind of documentation as to why the smaller aneurisms are being done.

But the SVS has a concern that if you're going to do smaller aneurisms where the indications are already a little shakier than for doing large aneurisms where the risk of rupture is much more real that you better have good outcomes and that's why those chose to do these separately as a marker, you know, for quality.

Arden Morris: Okay, so you're seeing this primarily as a separate measure and not a competing measure because they're talking about small aneurisms?

Vivienne Halpern: Yes, because none of the other ones sort those out and they particularly addressed size because if you're going to do it for a smaller aneurism, you're really - not that you shouldn't have good outcomes for any size - but you know, if you're doing a seven-sonometer aneurism, you know, the patient has a bad, you know, a 30% risk of rupturing.

So as opposed to somebody with a 5.5 or less sonometer aneurism, if it's less than 5.5, at that size they have a five to 7% chance of rupturing so your outcome - your risk of surgery - has to be less than that and the way they do it was actually kind of clever because, you know, they want people to make sure that they're looking at the rest of the patient basically.

So if you're going to do a five-sonometer aneurism, it shouldn't be in a guy who's, you know, like 70 on home oxygen and an EF of 25% because the chances of them doing well or living long is not good even sometimes with an endovascular.

Arden Morris: Okay.

Vivienne Halpern: So if you're going to do it, you better be doing it for the right reasons and have a good outcome so I see them as different.

Arden Morris: Okay. Do you see any downsides to measure 1523 or 1534?

Vivienne Halpern: Well, the issue again always comes up with the registry issue because you do need some kind of registry. It doesn't have to be the SVS registry but you do need some kind of registry where the size will be documented.

Now everybody should be documenting the size of the aneurism in the chart before they fix it because that is the main indication of size so it shouldn't be a hard piece of data to extract.

Arden Morris: So when you say that the issue arises again of the registry meaning that the hospitals...

Vivienne Halpern: Well, because we keep having this discussion of registries and is one registry better than another.

Arden Morris: Belongs to a registry, right.

Vivienne Halpern: Yes. Now the SVS has I think gotten through. I don't know if there's anybody on the SVS from the call so maybe they can answer this but I think they were trying to put through a code to indicate size and I think that it got accepted. It was the last thing that I heard.

Arden Morris: You mean a CMS code?

Vivienne Halpern: Yes, so that would make it easier to pick up from a just claims data point of view too.

Arden Morris: Is there anybody on the line from SVS?

Female: Are you asking STS or SVS?

Vivienne Halpern: SV...

Arden Morris: I thought this was SVS, Society of Vascular Surgeons, isn't it?

Vivienne Halpern: Yes.

Arden Morris: It's not the Society of Thoracic Surgeons, is it?

Vivienne Halpern: No, it's the Society for Vascular Surgeons.

Arden Morris: Yes.

(Heidi): This is (Heidi). I think Tim Kresowik's on the webinar. I don't know if he's - is his line open - I know he's the one SVS representative we have on the - is Tim around?

Tim Kresowik: I am but I'm not - my line must be muted somehow.

(Heidi): No, we can hear you now.

Tim Kresowik: Okay, sorry, so let me just explain this rationale of what we're trying to do here and Vivienne has done a nice job but the thing that we're trying to counteract which if you look from a health policy point of view, one of our biggest problems we're all facing is cost.

And I would have to agree with what Vivienne is saying is that right now there's especially in the endovascular arena there's probably tremendous overuse of endovascular aneurism repair for aneurisms on patients that would be best having nothing done.

So our purpose here is to try to create a measure that in (offense) is self-risk adjustment so you don't need risk adjustment and in fact by risk adjusting you would be rewarding people for doing the wrong thing.

So the principle here is that if you are doing someone that has a moderate or small aneurism, you should factor-in to your decision-making their age, their expected longevity, their co-morbidities because you're doing a purely prophylactic procedure whose only benefit can be to perhaps prevent rupture.

So if you're doing it on a small enough aneurism that's unlikely to rupture in the next five years on someone who has a life expectancy of two years, you've obviously not benefited anyone so that's the basic principle and I do want to comment related to the administrative versus registry data.

The only reason this is coming before you as a registry is because that's the only avenue we had for bringing it forward. We do have now and it's gone through the CPT-2 approval process is to have administrative codes that people that people can size aneurisms and present this data.

Now the beauty of this approach too is that it becomes self if you will promoting in that if down the road once these codes are available, if people don't report these codes if you assume that they fall into the small or moderate category, you again have the ability to capture 100% of patients being done so it will be eventually available through an administrative route.

But because of the rules at NQF, we could only bring it forth in this manner so my major questions that go forward in terms of talking about harmonization is, you know, at what point will these be rediscussed and, you know, if they're going to be seen as competing, you know, how will that be resolved in the future so I don't know how that will impact on this discussion now.

But I do think in relationship to things like volume which actually has a very perverse incentive to increase the overutilization that's happening out there now so in fact from a health policy point of view it's probably destructive to have volume as an indicator of quality because many of these institutions are if anything overutilizing as it is now.

Arden Morris: Okay, so if we stick with the SVS-submitted measures which don't include volume and are really about small AAAs, I agree with Vivienne that this is not a competing measure based on what you've said so far. Anybody disagree with that? Anybody think that this is a competing measure with the AHRQ measures?

And then the second point about it being registry data, I think that that's, you know, as was pointed out that's basically what we have right now although in the future we'll have something different and we'll be able to potentially use administrative data for these purposes when the CPT-2 codes are widely used.

Alexis, would it be appropriate for us to say that in the future that we anticipate that this could potentially be harmonized with the other measure?

Alexis Forman-Morgan: Sure. (Heidi), did you want to add to that?

(Heidi): Yes, I think so just to add-on to what Tim said, because they did not have the testing on the use of the CPT Category 2 codes because they didn't have the codes yet.

Alexis Forman-Morgan: Right.

(Heidi): That was one reason why we couldn't bring them forward but I think when they do bring it forward, it will either result in depending on the timing an ad hoc review where we will look to see how the measure has fared using those codes.

And then at the next review again where we would be looking at the AHRQ measures, potentially the Leapfrog measure depending on the timing as well as the SVS measures assuming they all go forward.

If they bring forward the claims data then I think the expectation would be Number 1 that they are harmonized to the extent they are. They do measure different things. I think we've talked that through and it's very at least clear to me the differences.

And then look to see that that harmonization has occurred because I think unless for some reason there's a different interpretation when there's the next review, that would be what we would be looking for and that's the expectation that I think you all can put forward now but if that does happen, you would want to see that.

Arden Morris: Okay, I'm all for putting that expectation forward. Does anybody disagree? Anybody want to speak up, ask any questions or raise any conflicts with this?

Male: Sounds good.

Arden Morris: All right and Rick anything else that you want to add in terms of like thinking about elderly people and their lifespan in terms and providing best care? I remember you had some comments previously, some time ago.

Richard Dutton: No, I think these measures are okay as we've discussed. I'm in agreement with what's been said.

Arden Morris: Okay, all right, I think we can probably move on from that then. The next item on the agenda is...

Male: Thank you guys.

Arden Morris: Thank you. Let's see, I need to find my agenda here with legible font. Is it time for us to move on to harmonization of prophylactic antibiotic measures?

Alexis Forman-Morgan: Yes.

Arden Morris: All right, great, we're right on time which is nice. Let's see, so we asked the developers to provide plans on harmonizer and combining multiple prophylactic antibiotic measures and that won't be available for some time.

We wanted for the developers to provide a plan to harmonize to the committee and Alexis included their response in our document. They haven't had time to produce this plan. NQF, see I've recommended that we can't wait for harmonization so we could consider recommending both for endorsements or simply discuss a little further before we get to our recommendations.

Alexis Forman-Morgan: Right and this is Alexis. Just to add to that so for this first group we have two maintenance measures and the discontinued group 529 and 128 and so the committee held off on voting on final recommendations for endorsement because they had these suggestions for harmonization regarding the exclusions and the end times.

And we did take this back to the various measure developers and they are currently working to schedule a conference call to begin discussions on how to harmonize and unfortunately it won't happen most likely within this project.

So what staff is suggesting which is similar to the next few measures and which you've already voted on, if you do like the measures as they stand currently, you can vote for the measures with

the expectation that those measures will be harmonized and submitted to the next surgery project which is scheduled to launch in 2013.

Male: Do we have another option?

Alexis Forman-Morgan: Yes, you can vote the measures down as they currently stand.

Male: Down, right, yes.

Male: Practical status is that they'll have to coexist, correct, at least?

Female: What is the downside to them coexisting?

Dale Bratzler: This is Dale Bratzler, one of the developers. You know, these measures, some are used in the hospital setting and the others are the PCI measures which are collected through a very different mechanism.

Male: So until they're harmonized, we shouldn't be voting them down, shouldn't be closing them then, should we?

Arden Morris: If we did vote them down, would that mean that they would not be harmonized in the future?

Dale Bratzler: Well, it's not that easy.

(Heidi): Yes, this is (Heidi). I mean, if you remove endorsement, there's no reason to not expect that they would continue working on harmonization and I think they would have the opportunity to bring

back the measures for consideration at the next opportunity with the expectation that the harmonization would be done by then.

But you have to weigh the implications and it truly is your call on whether having the measures there while not perfect yet or harmonized yet is a better value at the moment, than having measures - the endorsement removed - and waiting until the next time we know for sure would be 2013 where they could bring the measures back in and so I truly it is your call as to how you would like to proceed on that.

Arden Morris: Does anybody see a downside to endorsing the measure?

Male: I don't.

Male: I don't.

Female: I don't.

Arden Morris: How about you Carol?

Carol Wilhoit: I agree.

Arden Morris: Okay. Well, I think that we should proceed then. I think we should go ahead and recommend it. We endorse and, you know, go with the staff recommendation that these could be submitted with the expectation that they would be harmonized in the future although it's not that easy as Dale said. Anybody disagree?

Male: No.

Male: No.

Arden Morris: Great, okay, so let's see here. Let me pull-up my agenda again. Anything else that we need to discuss with regard to those measures, Alexis?

Alexis Forman-Morgan: No. I just wanted to point out the second-set selection of the antibiotic measures and then that third set, timing and received. The committee has already voted on the maintenance measures and you did vote with the expectation that those measures will be combined with the currently-endorsed measures and submitted under the next project in 2013.

But you did ask for a harmonization plan and so I just wanted to provide the measure available response and it's the same response from the first set. They're currently working together to figure out how to combine and harmonize the measures but I just wanted to provide that information to you all but there's nothing that the committee needs to vote on, no action taken.

Arden Morris: Okay. Our next agenda item is a review of measure 1741, the consumer assessment of healthcare providers and systems, the CAHPS surgical care survey and I do not believe that this was assigned to any of the committee members to introduce.

It replaced another measure. It was submitted by the American College of Surgeons and we wanted to have a representative from the American College of Surgeons give the introduction for the measure. Is there somebody on the line right now that can do that?

Alexis Forman-Morgan: Joann from Westat should be giving a brief overview. Joann, are you on the line?

Arden Morris: She's on the Website.

Operator: Well then, Joann, your line's open.

Joann Sorra: Okay, hello. Can you hear me?

Arden Morris: Yes.

Joann Sorra: Okay, well thank you all for this time and I'll just give a really brief overview and then if we can go into questions and answers but the CAHPS surgical survey is basically a patient experience survey and set of measures that's designed to specifically assess patients' experiences with surgical care.

And it's based from and informed by the existing CAHPS instruments which assess patient experience in other settings of care like the hospitals and the outpatient ambulatory settings with the clinician in-group survey.

And the ACS basically funded the development of this survey in conjunction with the Surgical Quality Alliance and its basic goal is to measure aspects of surgical quality that are important to consumers and for which consumers are actually the best source of information.

So it measures things such as how well the surgeon communicates with patients before surgery, surgeons' attentiveness on the day of surgery, information that patients receive to help them recover from surgery.

So what we are proposing is that the six composites that are included in the survey are up for endorsement plus a single item global rating which asks about these, basically an overall rating of the doctors, the surgeons.

So that's pretty much the very quick overview and if we want more of a background, then I can also go into the development process and sort of the fundamental information that's in the package.

Arden Morris: Can you talk about how it was validated?

Joann Sorra: Sure, so as I said the surgical survey was informed by the existing capture data as well as a literature review that was conducted by the American Institute for Research under contract with ACS and so a literature review was done. There were two rounds of focus groups with consumers asking them questions about what they consider to be good quality surgical care.

There were items - there was an item development process - that was then proceeded through cognitive testing where the surveys were - the survey items - were completed by patients and then patients responded about how well they understood the questions and whether they thought the questions were measuring important aspects of surgical care.

And then a field test was employed with patients - and let me find the information about that - the field test was done in the summer of 2008 and included patients from 96 surgeons, 33 practices across nine surgical specialties.

And so we received - AIR received - data from a total of 2750 patients and that data was then psychometrically analyzed to look at the factor structure of the items to see how well they held together as concepts to assess their reliability and that information is included in the packet.

After the psychometric - after the field tests - and psychometric results, the CAHPS consortium and the CAHPS consortium was involved in the process with ACS and AIR throughout and when the consortium reviewed the results from the psychometric analysis, they asked for some additional testing on the labels that were used for the composite and that work was done.

And then some modification was done to the composites and the titles of the composites so the ones that we are showing you here in the submission and providing for your consideration are the final composites so there are I think said composites and then one full rating and the composites have anywhere from two to four items.

Arden Morris: Okay.

Alexis Forman-Morgan: This is Alexis from NQF and I just wanted to let the committee know that we did do a staff evaluation of the measure and we have Karen Pace on the call as well. She's our in-house methodologist. I'm not sure Karen if you wanted to make any statements before the committee began their evaluation.

Karen Pace: Yes, thanks. This is Karen Pace. I just wanted to clarify for everyone because it's a source of confusion for all of the CAHPS is that NQF endorses performance measures that are based on using the CAHPS survey as the data source so it's a little bit of a fine distinction but we're not simply endorsing the survey.

We're endorsing if, you know, if it's recommended and goes this way, NQF would be endorsing the performance measures that use the CAHPS survey as the data source and so I think it would be important to understand what are the performance measures and perhaps we can just get some clarification.

My understanding from reading this is that for each one of those measures - the composites or the global rating - that the developer is suggesting that the performance measure that for example would end-up getting reported if this were picked-up for public reporting would be the percentage of responses for each response category.

Or my question is because I know for some of the CAHPS and public reporting what's reported as the performance measure is just the top response category so maybe the measure developer could walk through what are the measures and how those are computed.

Joann Sorra: Certain rule. There are different ways that you can report so like you're saying you can report for example in the first composite called information to help you prepare for surgery, there are two items and the response items are yes, definitely; yes, somewhat; or no.

So the questions are sometime before your surgery, did anyone in the surgeon's office give you easy-to-understand instructions about getting ready for your surgery: yes, definitely; yes, somewhat; or no.

You could report the percentages of patient response to each of those response options or as you were saying Karen, you can report the clinical top box or only the percentage who answered yes, definitely.

Alternatively you could report a mean score but most people are not reporting mean scores because they have less sort of it's hard to understand what a mean is between a one, two and a three.

So what you see in the submission is that basically it's the percentage who are reporting the top five is what we're saying that probably the most often-reported - Carla, Carla Zema is also here, part of the team who is helping us prepare this submission as well as (Sophia Tucker-Cletus) who's here at Westat.

So I'm going to extend the invitation to Carla and (Sophia) to jump in at any time if they want to elaborate on any of my answers.

Female: Can I just clarify something with you, so suppose that 20% of the response are yes, definitely and a very small proportion like 5% is no and everything else is yes, somewhat for Hospital A and then for Hospital B, they have 30% of their responses as yes, definitely and they have 50% of their responses as no, then they're going to come out looking better, right, because they have more yes, definitelies? Is that correct?

Joann Sorra: Yes, in the top-box scoring approach, yes.

Female: Right, so but if you took the mean, then they would definitively look worse, right?

Joann Sorra: Yes.

Female: So but your decision was to go with the top-box scoring rather than the mean?

Joann Sorra: Yes, I mean, there are different - for different purposes - you can calculate the scores in different ways but for public reporting and particularly given that CMS reports is way over the hospital CAHPS, it's most like going to be the top box score.

Female: Okay.

Joann Sorra: And I have to say that you would almost never - you really almost never - see huge percentages in the negative range on any of the CAHPS surveys. It would be - I don't think I've ever - seen that where you have even as much as 50% in there. Carla, would you agree?

Carla Zema: I do and I think the other rationale for doing top-box scoring is to set the bar from an improvement perspective to say that the most important outcome is the top box.

Female: Well, I like that rationale. I think that that's a good rationale. On the other hand, it'd be really nice to know how well the mean correlates with top-box scoring in this particular, you know, in using the CAHPS surveys. Has anybody looked at how well it correlates?

Female: Anecdotally when we've had users and this is not specific to this survey for any CAHPS survey you can use mean scoring. You can use top-box scoring. The other advantage to top-box scoring is it's easier to understand from a consumer perspective versus the mean scoring.

Anecdotally we know users that have looked at their data both ways just to kind of see what the difference would be have not found that there are significant differences in how in essence the different units are compared when you do it one versus the other. I will clarify...

Joann Sorra: It sounds like we really don't know that nobody has really sort of compared these things and seen what the correlation is so we're sort of going with the probably.

Female: I did want to clarify that this is an ambulatory instrument mainly. Primarily it is sampled at the ambulatory level, at the surgeon level, or the surgical practice level so you had said Hospital A and Hospital B. I just wanted to clarify that we are talking about a survey that is administered based on the ambulatory setting.

Arden Morris: When you say ambulatory setting...

Female: Clinical?

Arden Morris: ...are there surgeries only ambulatory?

Female: No, no, no. The surgery could be an inpatient but it's based upon records that are based on a surgeon's ambulatory practice to we're talking about surgeons at their practice level and not necessarily out of hospital data.

Arden Morris: Okay, so at the clinic level and who would be responsible for implementing this? Would the surgeons bear the burden of or would the surgical clinics bear the burden of implementing and reporting?

Female: I'm going to take a stab and then I'll kind of turn it over to ACS as well for their response from an ACS and SQA perspective. From the CAHPS perspective, we see a number of different scenarios if you will of implementation. There are implementations at the community level where you do have an external organization that sponsors the implementation of it.

We also do see plenty of implementation of surveys like this that are implemented by the practices themselves not only for internal quality improvement but also now for example ABMS from the maintenance of certification requires some communication items that are based on the CAHPS clinician and group survey and they have also agreed to accept the surgical care survey instead for specialties for which this survey is more relevant.

Bob Jasik: This is Bob Jasik from the American College of Surgeons and I would just agree with all of that and basically just to say that you could use it the same way as the clinician and group and one of the reasons that we created this was because the questions just didn't fit right in that for surgeons for this...

Karen Pace: So this is Karen Pace again, so most of the information you've provided is about the item level analysis and I think you said you had something like 96 surgeons or practices in your testing pool; is that correct?

Joann Sorra: Yes.

Karen Pace: Can you provide any of the surgeon-level performance data without identifiers but to give the committee an idea of how these are scoring at the performance-measure level?

Joann Sorra: One of the appendices - I'm trying to find it here - Attachment I has the surgeon-level results from the surgical care field test and that basically shows you the - let me find that appendix - it shows you the minimum score for a surgeon, the maximum score which usually is 100% and then it shows the confidence interval and the mean for the surgeons.

Karen Pace: Oh, great.

Female: Do we have that?

Joann Sorra: It would be Appendix I...

Female: And where is that in the document we have?

Joann Sorra: That's in your appendix so no, it's not in the...

Female: It's listed at the end of the main document of the application and was provided as a temporary zip file.

Female: Right, so Alexis, maybe you can...

Alexis Forman-Morgan: Right, this is Alexis. This was sent out by Jessica when we sent the original measure submission form. There were a series number of attachments so we did not submit it or we did not attach it in these materials because of the document size so you all were sent it.

I can forward it to everyone very quickly again while we're on the call but it was sent on October 26th and I'll just forward it again.

Female: And Jessica can you project it too? I don't know if you can. That might be helpful for those on the webinar.

Jessica Weber: Sure, I can project it.

Female: Okay, we'll project it too. Let me help.

Karen Pace: I have a question about the data. I'm on Page 118 of the long document that we got which is the scores for all measure components and the scores are all really, really high. Are these top two box or are these top one box? What are these scores that we have in the document?

Joann Sorra: Can you tell me the name of the document?

Karen Pace: Well, the document is the one that...

(Crosstalk)

Karen Pace: ...the one they sent out for the meeting today that has the...

Joann Sorra: The submission form, you mean.

Karen Pace: ...the submission form and it's on Page 14 and 15 and 16 of the submission form.

Joann Sorra: Oh, okay.

Karen Pace: And so like for question 3 the score was 97.6%. For question 4 it was 98.9. The scores are really high and I'm just trying to understand and those it says positive response. Is positive response top one box, top two box?

Joann Sorra: The top box, yes.

Karen Pace: But right, I think that's a good question because and it's the same question about the Excel file because it says percent yes so that includes the yes, definitely and the yes, somewhat or just the yes, definitely?

Joann Sorra: That is yes so the field-test version of this instrument had yes/no items and for that very reason we so this is a survey that was developed - sponsored by ACS - developed by AIR, American Institute for Research outside of the CAHPS consortium. It got submitted to the CAHPS consortium for review and adoption into the CAHPS family of surveys.

When the CAHPS consortium reviewed the survey, we agreed that there was not enough variation given a standard two-point yes/no scale and so we required that they adopt the three-point yes/no scale which shows more variation so you will not see such severe ceiling effects.

Karen Pace: But it has the final version has a three box rather than like a five-box option?

Carla Zema: Correct, and the reason that it does not have - we don't use the standard for example a four-point frequency scale - is because the items are asking about specific things that happen or they don't or they tend to be one-time. If you only had one pre-operative visit, you can ask about a frequency.

Joann Sorra: And actually Carla I think in the field test there were a couple of versions again because AIR conducted it. It's our understanding of the exact details but there are some frequencies that do show a yes, definitely, yes, somewhat and no, somewhat and definitely no.

So there were some options that were tried in the field test with regards to the yes/no, yes, definitely, yes, somewhat, no, and then a four-point version actually.

Carla Zema: Right, but when they're reporting percentage yes, they are mapping the definitely yes and somewhat yes to their yes so in essence it is if it was a three-point scale, they are reporting two levels of the scale.

Joann Sorra: Yes, so if you read the note at the top of that document, it says it's kind of interpreting the scale so yes is yes for a simple yes/no or definitely yes, somewhat yes were ones that used that option so there were two options that were used, a straight yes/no and a definitely yes, somewhat yes.

And this is combining either just a yes or a definitely yes, somewhat yes but that was out of a four-point yes/no scale.

Carla Zema: So and just to clarify then, you don't have actual performance results on the final scale that you've submitted for NQF endorsement; is that correct or...

Joann Sorra: Well, the percentages are reporting where we have - where they used a yes/no - it's reporting the yes. Where they used a definitely yes, somewhat yes, a definitely no, somewhat no, definitely no, it's reporting the top two, definitely yes and somewhat yes.

So it's not exactly the same response option that's shown but as you can see it is true that they are very as we would think negatively viewed or very positive overall and that's not unusual for CAHPS items either.

Arden Morris: Okay, so in terms of the issues with this measure so far, and I have - this is Arden - and I have to say that I think it looks like a very well thought out survey. We really were very strongly in favor of measures around patient centeredness in our in-person meeting and this would absolutely fit that.

There's still some issues that are outstanding or unclear to me. One is that we don't know what the final scale responses look like. We don't have that information. A second outstanding issue is that we really don't know what the correlation is between the mean score and the top-box score.

And it would be nice to know that to have some kind of reassurance that using the top-box score is the right approach to this and then thirdly it's still not really clear to me who bears the cost of conducting the survey and reporting the results.

It's not clear to me whether that would be a cost that would fall upon individual clinics or if it would fall on some external body or exactly where it would fall so those things remain unclear. Are there other issues that are unclear for the steering committee members? Anybody on the steering committee that would like to bring up any other issues?

Karen Pace: I would respond to one thing about cost and, you know, and who does the survey. With the other CAHPS surveys, with the health plan survey, with the (H-CAHPS), with the clinician and group survey, therein the surveys are in the public domain and anybody who wants to do it can do it and anybody who wants to do the survey bears the cost.

And so I think that, you know, the answer in terms of cost is probably it depends on who wants the results and why they want them because it could be, you know, it could be done by a health plan. It could be done by a hospital. It could be done by physicians themselves.

I think it would depend upon, you know, the particular setting. I don't think the fact that there's a survey doesn't mean that there's just one way of paying for it.

Arden Morris: Okay, so this would potentially not really have anything to do with CMS and would not necessarily be required for providers to submit these data?

Karen Pace: Right, and if CMS did require it, then for that purpose they might mandate that the physicians pay for it for example but the survey itself, a CAHPS survey is typically in the public domain and so if we as a health plan want to use it, we can do it.

We can pay for it. We can look at the results and, you know, or CMS might at some point mandate it. That would be up to CMS but it would, you know, I think there's - the survey itself - is independent of the payment mechanism.

Arden Morris: One of the things that we have looked at as we're reviewing these quality measures is what the burden is on for most of them on hospitals which wouldn't apply to this particular situation but that has been part of our discussion. Does it not belong in our discussion with this current measure?

Female: Yes, because if it's just to do surveys, then why does it need to be endorsed?

Karen Pace: Well if say we as a health plan wanted data on surgeons and wanted to get survey information and we chose to pay for it, we would rather use - we would rather not start - from scratch and come up with our own survey and maybe people like it, maybe people don't.

Maybe people think we have yes/no and we should have had five options, you know, so if we wanted to do a survey, it'd be really nice to have an endorsed nationally-standardized survey that we could go to and it's in the public domain and we could use it and it has some face credibility.

It has comparison data available and so that that would be a really good reason or as a practice, you know, say you've got a group of 20 surgeons and they wonder how they're doing and what their patients think of them and they wanted to do it for their own feedback and quality improvement purposes and not ever tell anybody the results but just use it for their own benefit, again they would have something with some comparison data.

Because if you just survey your own patients and your results are 90%, well is 90% a good rate or a bad rate? Well, if you've got comparison data and the national average is 80 or the national average is 95, that helps you interpret your own 90 so I think that's the reason that the endorsement is helpful because it could be done in a variety of settings.

Bob Jasik: This is Bob at ACS and just to add to that too, we've had conversations with CMS and we're trying to promote the use of this with them and I think the NQF endorsement would be really helpful in that and they've actually already proposed to incorporate the surgical CAHPS into the OPSS quality reporting program in the ambulatory surgery center quality program as well.

Arden Morris: Thank you. Am I the only one on the committee that's concerned about the fact that we're not really seeing the final scale and we don't know the correlation between the mean and the top-box scores?

(Anne): No, I'm a little concerned about that too. This is (Anne).

Karen Pace: And I'd certainly like to see what the results are and get a sense and see a distribution. If everybody's got the same score, it's not particularly useful so that's where having some real data helps.

Arden Morris: Would it be possible for the developer to get back to us with that data?

Joann Sorra: So the document that we were looking at that had the percent yes on it provides some information about variability so if we go back to that document, you can see that the minimum score here on some of the items is...

Karen Pace: Which document are you talking about? Are you talking about the submission documents that we have in today's meeting or one of the competitive...

Female: No, she's in Appendix I.

Joann Sorra: Appendix I.

Karen Pace: Okay.

Joann Sorra: This is the one that has surgical character of A results from field tests so it has the items in the field test, not all of which are in the final but this is just to again these were all of the items that were included in the field test and then it was narrowed down and as we said, the field test tried different numbers of response options for yes/no.

And what the final response options are which are standard on the clinical and group CAHPS survey are yes, definitely, yes, somewhat and no and those final options are on Page 17 of the submission so we have finalized the response options to the three-point frequency.

There is one set of questions that is just a yes/no and that's on surgeons' attentiveness on the day of surgery at the top of Page 18 and then the overall rating is a zero to 10 which is consistent with the other top surveys.

Karen Pace: This is Karen again so what you've given us on these results is item by item but can you at least put together what the score would be based on what you're suggesting goes into a composite because you're not suggesting scoring item by item. These are going to be combined into domains, correct?

Joann Sorra: So there's a little of both. There's reporting at the item level as well as at the composite level. The difficulty with generating a score at the composite level with the field test data is that some of the wording was slightly modified from the field test to the final.

And as I just indicated, some of the items in the field test used different frequency scales that we would be combining the definitely yes and the somewhat yes on some items with the straight yes/no on other items so it still really doesn't get you to the final version which is informed by the cyclometrics, by follow-up focus groups on the composites themselves.

And so really there isn't a lot of data on the final instrument as we are proposing it because this is a shortened version and it's not exactly what was field-tested so it would be difficult to and not really informative of the ultimate final version for us to take the field test data and combine the items because they use these different response scales.

So really the best information is at the item level, at least to see the variability of response and the general percent that the what is roughly equivalent across the items percent yes.

Karen Pace: And is there any work going on right now in terms of further testing of this final or what you think is final CAHPS and the measures or has that ended?

Joann Sorra: Carla and Jill are probably aware of some entities now that are using the instrument and I don't know if you have the details Carla or (Joe) on that?

Carla Zema: This is Carla. I was actually looking at some of the data that came out of the AAO - it's the otolaryngologists - did a feasibility study of it and they did go ahead and calculate the composite.

I believe that they used mean-level scoring and so I'm seeing a range of 66% through 90, high 90s percentages in terms of the mean scoring and then in terms of the top-box results that they found again ranging from about 80% through a high 90% at various sites.

And that is using the final instrument but none of these implementations are sponsored by the CAHPS consortium so we aren't truly owners of the data because the CAHPS consortium doesn't implement the surveys. Again, we are the public domain, you know, we put the survey in the public domain.

Arden Morris: Right. I'd say as a group, as a committee, if we wind-up voting to endorse it, part of our endorsement should include how it should be scored because if we're trying to basically endorse something that could be widely used by whomever chooses to use it because it's been validated, it's reliable and for comparability, then part of that should include how it ultimate is scored and compared.

It's not going to be comparable if some people are using top box, some people are using top two boxes, some people are using mean.

Karen Pace: Right and this is Karen and that's absolutely right in terms of a performance measure, a national voluntary consensus standard, it has to be something that's going to be used consistently and what they've submitted for specifications is using the top box I believe.

Carla Zema: And you have a document that's CPT codes where they've only tried out for those CPT codes.

Joann Sorra: I'm sorry; could you ask that question again?

Carla Zema: It is that one of the attached documents, the last one, it's CPT codes for surgical CAHPS 90-day globals. Was it only tested on those CPT codes?

Joann Sorra: Those are the CPT codes that each surgical specialty felt represented the surgeries that they felt should be included.

Carla Zema: Because it's kind of an odd array. There's only a couple. I only see a couple of them on here and actually most of it is vascular surgery but one is excision of external ear and there's not that many CPT codes. That's why I thought they were kind of and it's a kind of just a weird array.

Joann Sorra: If you look at the row numbers, there are actually 17,000 CPT codes.

Carla Zema: Okay, so they just showed a small section of them here.

Joann Sorra: Correct.

Carla Zema: Okay.

Joann Sorra: But the row numbers are the actual row numbers so there are over 17,000 CPT codes that were identified across numerous specialties.

Carla Zema: Okay.

Joann Sorra: Just going back to the scoring, I just wanted to clarify that in the field test, different scoring methods and response options were tested but in the final recommendation on Page 17, the final scale is yes, definitely, yes, somewhat, no with yes, definitely being the top-box proportion and the score that would be reported so we're not recommending mean scoring essentially.

Researchers may want to do mean scoring but again, we're recommending top-box scoring.

Arden Morris: Anybody else from the steering committee with questions? Anybody else from the committee who has questions about this?

Terry Rogers: Can I admit to a little confusion?

Arden Morris: Is that Terry?

Terry Rogers: Yes. Yes, it's fine. I don't want it to slow down the process. It seems like - I've been reviewing some of these documents - they're very lengthy and so there's a lot of stuff there.

Arden Morris: Yes, well this is an incredibly big project. It's very complex and I think that it was obviously undertaken with a lot of rigor.

Terry Rogers: Yes, yes.

Arden Morris: So I think it's a terrific survey and I think it serves a great purpose. I just am hoping that if we vote to endorse it that it will be, you know, that it won't get - that the process won't break down - because of confusion about scoring or the meaning of the scores.

Terry Rogers: Right, so could we be very directive and prescriptive about that because I agree with you.

It looks like a lot of rigor in fact was addressed and sent. Addressing what should be fundamentally a fairly simple question, did your medical care work and were you happy with it?

And I agree that without having some sort of common answering matrix, that is, if everybody has to speaking the same language because what's going to happen is our scores say we're the best. Well, if someone else is using another kind of measurement of grids, then best against whom?

You're setting-up for apples and oranges as I'm sure you all know so is there some way that we as a committee can say we're inclined to endorse this but only if you can prescribe a common method of reporting results.

Joann Sorra: And I'm sorry, this is Joann. I do want to clarify that that prescription is actually already out there and available and extremely documented but again these are in the attachments as maybe you weren't looking through.

So Attachment G has instructions for analyzing data from CAHPS surveys and Attachment H has the reporting measures for the CAHPS surgical care survey and that provides specific instructions for how to report the overall ratings using the top-box scoring method.

Terry Rogers: Okay, so then why are we talking about variability in reporting?

Arden Morris: Because there was confusion around it. It wasn't clear from the documentation that we were looking at that that was actually the process and because the measure developers as they introduced it talked about different methods for reporting and it was really unclear.

Also the data that we received was put forward in a variety of ways. It was only yes, only no for some of them. It was the yes, definitelies and yes, somewhats were combined for some of them.

So it's right that you're confused and that, you know, we as a steering committee are paying attention to this issue although the measure developers say that they have stipulated it clearly in their attachment, obviously it needs to be more clear because we're spending a whole lot of time discussing it.

Female: I think one of the things that makes this process unclear is that as the CAHPS consortium we have not made a firm recommendation in terms of scoring. We've kind of said you have options. There are philosophical reasons. There are many reasons why a survey sponsor might choose to do different types of scoring.

And so in testing of the CAHPS surveys, we often look at the scores in a variety of ways to make sure that the psychometrics hold up depending on how you look at the data and so what you're seeing is that variety of all the different ways that we look at the data.

The reason that we have recommended top-box scoring is because we are seeing a trend in the industry now with clinician and group and the hospital survey that public reporting is moving towards top-box scoring.

So we are making a definitive recommendation to you despite the numerous ways that we've looked at the data that for consistency both in how people are used to seeing CAHPS data as well as what we anticipate future sponsors of the survey might require, we are recommending the top-box scoring to you.

Arden Morris: Okay, and that's what we need because we're not a research group. We're a policy group.

Female: Correct.

Arden Morris: And we're talking about policy which is a little bit different and I think that this - I think it needs to be - extremely clear. Is there anybody else on the steering committee that would like to say anything else about this, about this measure before we move on to start talking about importance and the rest of the criteria?

Vivienne Halpern: I would just say I just scanned through the documents to say how to analyze and it was still confusing so I agree that it needs to be clarified exactly how people want this data to be analyzed. Otherwise, I think it will because even if you have that document, they still give you a variety of choices.

Arden Morris: So Karen and Alexis and anybody else and (Heidi) maybe, would it be appropriate if this documentation - if this measure - is endorsed and this documentation is made available, would it be appropriate for us to ask for a change in the documentation so that it doesn't look as much like opportunities for research as it does like a very clear sort of policy endorsement?

Karen Pace: This is Karen and I think that can be probably clarified some more in the actual measure specification so that it's very clear what is up for NQF endorsement versus obviously all of the other documentation follows a lot of the format that's been setup by the CAHPS consortium.

But we can certainly look at the specifications and make sure that those are clear in terms of how the measure that's being put forward for endorsement would be computed.

Arden Morris: Thank you. Anybody else on the committee with any other comments? Questions?

Karen Pace: One question I have is, I mean, I think from the discussion everybody likes the idea of this but at the same time I think there's a certain amount of discomfort with the fact that final data using the final methodology and so on, you know, results aren't really available.

Is it feasible to endorse this as a time-limited measure or something like that to sort of promote to get it going but at the same time to recognize it's not quite final and there's still a little bit of flux involved?

Arden Morris: And we're looking for an answer from the NQF on that?

(Heidi): Yes, this is (Heidi). I mean, unfortunately this type of measure would not really classify within the criteria for time limited so I don't think we can do that. We need to - thinking through next steps - we can work with the developer like Karen said to make sure that it's more clear.

But at the end of the day, if you continue to have concerns that the measure once we make the information more clear still does not meet the criteria and is not yet ready for public reporting, then I think you need to consider that and perhaps not recommend it.

At this point it's hard to do much more than that. It all depends on what the concerns are and how quickly the developer could provide that information back to us but there is no opportunity for time limited unfortunately.

Arden Morris: Then I would say since we have a lot of concerns, should we just be saying that we'd like to revisit this one, not vote on it right now?

Vivienne Halpern: So (Heidi), I don't know, you can answer that question but I just also what is the timing of the patience experience with care project?

(Heidi): Yes, hold on. I have it right here, so we do have one additional opportunity. We were trying to see if this measure was - you did feel it was ready - then we wanted to not wait for the upcoming project that Karen just mentioned but there is the opportunity for them to bring it forward and it is I believe submission is in July of this next year 2012.

So again it does mean that this measure would not be endorsed until closer to the end of 2012. If we did wait, that would be the one downside but there is an opportunity next year to have this reviewed.

Vivienne Halpern: Well, this would be up to (Heidi) and Alexis but I mean, you may want to go ahead and have the committee vote so that you really truly know the full range of the committee's thinking about this before, you know, you would decide to postpone or go one way or the other.

(Heidi): I agree. I think that's a good idea if everyone's willing.

Arden Morris: So are we voting on the SurveyMonkey because not everybody's on the call?

Vivienne Halpern: Yes.

(Heidi): All those will happen via SurveyMonkey.

Carla Zema: I'm sorry, this is Carla. Can I just ask one clarifying question? I know the basis of the concern is that you don't have data from the final survey and I'm going to kind of ask this in conjunction with ACS as well is would it be possible for us to get permission to obtain data from those that we know have implemented the survey to see if they're willing to share their data?

We don't own the data so we didn't put it in the submission but that sounds like that might address the concerns and that could potentially be turned around fairly quickly to you?

(Heidi): And I think that would address a lot of the concerns if that were available.

Carla Zema: And the reason that I'm asking also ACS because I think it's something that we could do together since it's a lot of their members.

Arden Morris: Alexis, what's our timeline on voting for this?

Alexis Forman-Morgan: The committee needs to submit their votes by next Friday, December 9th.

Arden Morris: Do you think that Carla do you think you could get us that information?

Carla Zema: I actually have a draft manuscript that they've worked on based on the data that they've submitted so like I said I know that it exists and I can look at the data. I just like I said I could not share that without their permission because it's not...

Arden Morris: I understand. Do you think that you would be able to get it to us by say Wednesday?

Carla Zema: I don't think that would be a problem.

Arden Morris: Okay. Well, let's give that a try.

(Heidi): Okay, so this is (Heidi), so we can wait - perhaps even it's up to Alexis how she'd like to do this - wait and circulate the SurveyMonkey on everything as well as the information if they're able to get it to us by Wednesday or Thursday.

And then I guess the one question I have for the group is do you feel that you would like to have another conversation about this measure by conference call or do you think you can handle the remaining issues once we get this to you by e-mail?

Female: E-mail's fine for me.

Arden Morris: I think that it would probably just be too cumbersome for the committee to have another conference call about it. When you say by Wednesday or Thursday, we don't mean tomorrow or the day after, right? We're talking about next week? I don't know if I'm not sure if we should receive it by Thursday, I don't know that that's really enough time to review it and then vote by Friday.

Carla Zema: I really think that I should be able to get it to you sooner than that. Like I said, I just needed to reach out to the AAO to get permission to share the data.

Arden Morris: I think the other question though was about clarifying. I don't know if there needs to be some sort of revamping of the submission clarifying how the data was going to be analyzed, right? Wasn't that the other question was what exactly are we using for data analysis?

Vivienne Halpern: Right, they changed the specifications to make it more clear.

Carla Zema: Well, we've tested very simple on Page 17 of the submission for this endorsement purpose but the larger CAHPS consortium again does have a lot of options and some very complex options for risk adjustment in all of this.

But users again some users use that, some don't so what we're recommending for this endorsement is on Page 17 which is simply top-box proportion and so I just want to be really clear about that, its purpose.

The other supporting documentation we've provided but it sounds like maybe that's confusing the matter more than helping the matter but...

Arden Morris: I think it is.

Carla Zema: ...but the submission itself on Page 17 is pretty clear and...

Arden Morris: What page is that in ours because our numbers are different?

Carla Zema: 120.

Female: It's Page 120.

Karen Pace: This is Karen and also what may be a little confusing is that in the I think in the section on scoring - some places it kind of talks about the percentage for each of the responses - so I think we just need to make sure that it's consistent throughout but we can go through and talk with the developer about that just to make sure that it's, you know, one way throughout the document.

Arden Morris: I think that would be helpful.

Female: Okay.

Arden Morris: So is that our plan? Do we need to do anything further with this discussion, Alexis?

Alexis Forman-Morgan: No, we will wait for ACS to submit the additional data and we will work with them to update the submission form to make things more clear.

Arden Morris: Okay. Let's see, then I think the only thing that we have left if there are no other steering committee comments is member and public comment.

Alexis Forman-Morgan: And Catherine, can you please open the lines for questions, please?

Operator: Absolutely. It looks like everyone is open except for Bob Jasik and that's it. Everyone else on the line, their lines are all open.

Arden Morris: Any member of public comments? Okay, Alexis would you like to move on to next steps?

Alexis Forman-Morgan: Sure, so once we receive the information from ACS, I will send that information out along with the updated submission form and I will send the link to the SurveyMonkey so that the committee can vote on the measures discussed on today's call.

Just an FYI, some of the measures including the AAA AHRQ measures, the committee will be voting on the four major criteria. During the May 4th and 5th meeting, the committee decided to hold off on voting on the four major criteria until the measure developer had updated the measures.

So you will be voting on importance, scientific acceptability, (disability), feasibility. You will be voting on those four major criteria, also for the surgical CAHPS measure that we just reviewed.

Some of the other measures as far as the SVS measures, those AAA measures, you'll do some voting on whether or not you recommend the measure for endorsement along with the prophylactic measures and I just wanted to give a brief update on the phase 1 measures.

Those measures were reviewed by (C-sac) and the measures were approved. They're currently with the board and we should have final ratification by the end of this week and then those measures will go out for a 30-day appeal process and I'll send out the press release once those measures have been ratified.

And the phase 2 measures are currently in voting and voting closes on December 5 and those measures will go to (C-sac) on December 12 and I will send out the final (C-sac) agenda for anyone who would like to call-in and listen to the discussion on the 12th.

I'll send out that agenda later in the week once it's been finalized. Are there any questions from the committee about anything that I just went over? Okay, (Heidi), do you have any additional comments?

(Heidi): No, other than I'd again thank the committee. This has been a phenomenal amount of work that we know you've done and it is truly very much appreciated by us and I just want you all to know that again.

Arden Morris: All right, I think that's it then.

Alexis Forman-Morgan: Okay.

Arden Morris: Well thanks, everybody.

Alexis Forman-Morgan: Thanks.

Female: Thank you. Bye bye.

Female: Thank you.

Female: Bye now.

END