

NATIONAL QUALITY FORUM

**Moderator: Ashlie Wilbon
May 2, 2011
2:00 pm CT**

Operator: Good day everyone and welcome to today's conference. All lines will be open for today's call.

We ask that you utilize your mute button when not speaking.

Also, today's call is being recorded. Please go ahead.

Ashlie Wilbon: Okay, thank you. This is Ashlie Wilbon from NQF. I just wanted to welcome everyone to today's call. Glad that everyone was able to make it.

Today's call if going to be focused on evaluating our very first resource use measure for this project - it's exciting. And we are going to be talking about the HealthPartners measure.

Just a quick review of the agenda. We did have a little rearranging from the last one we sent out. We just wanted to orient everyone to the plan for today.

We're going to start out in just a second with Ann Hammersmith who is our General Counsel here at NQF. She is going to lead everyone through roll-call and give some guidance on your disclosure ventures.

This is our first call reviewing measures so we will have to have everyone go around and disclose any potential or real conflicts of interest. And like I said she will give more instructions around that and how you should go about doing that as we enter in reviewing these measures.

Then we'll come back to NQF Staff, (Sally) and I will go over some of the goals and objectives for this call. That will be followed by a presentation from Sue Knudson who is at HealthPartners.

HealthPartners only submitted one measure to the project so while the other developers who submitted more than one measure presented during that General Message Webinar last week.

Where you're able to get an idea of how their general measures applied across multiple measures. Since they only had one measure we're giving them the opportunity to present during this webinar. And introduce their measure and queue it up for you so you can begin your discussion following that.

Like I said after their discussion, we'll hand it over to Bruce and Doris who are the co-chairs for the steering committee and they will lead the committee through a discussion of the measures at that time.

And we will open it up for public comment at the end of the, and adjourn by 5 o'clock. So let's go ahead and get started. I'm going to hand it over to Ann Hammersmith to do our roll call and DOI.

Ann Hammersmith: Hi everyone. This is Ann Hammersmith as Ashlie said, I am going to lead you through the introductions and the conflict of interest disclosure which we will do at the same time. I want to remind you of a few things before we start. We certainly don't expect to recount your CV to us.

What we would like you to do is to mention anything that you disclose that you feel is relevant to your service on this group. And that your fellow committee members should know about.

In particular we ask that you disclose if you have had any involvement with measures that compete with those before you today even if those measures are not under review in this project and have not been brought forward for endorsement.

I also want to remind you that you sit as individuals on this committee. You don't represent the organization for which you work. Nor do you represent the interest of any organization that may have nominated you to serve on the committee.

What I will do is I will call you name. If I mispronounce it, please correct me. I ask that you tell us who you are with and then do the disclosure.

I am going to start with Doris Lotz.

Doris Lotz: Yes, I'm Doris Lotz. I'm with the New Hampshire Department of Health and Human Services, as the Medicaid Medical Director. And I have nothing to disclose.

Ann Hammersmith: Okay. Thank you. Bruce Steinwald?

Bruce Steinwald: I'm an Independent Consultant in Health Economics in Washington, D.C. And I have nothing to disclose.

Ann Hammersmith: Okay. Paul Barnett.

Paul Barnett: I am the Director of the VA Health Economics Resource Center and I have no competing interests.

Ann Hammersmith: Okay. Jack Bowhan?

Jack Bowhan: Yes, Jack Bowhan, Wisconsin Collaborative for Healthcare Quality, nothing to disclose.

Ann Hammersmith: Okay. Jephta Curtis?

Is Jephta Curtis on the phone? Okay.

(Curtis Elword)? (Curtis Elword).

(William Golden)? (William Golden)?

Lisa Grabert?

Lisa Grabert: American Hospital Association and I don't have anything to disclose.

Ann Hammersmith: Okay, thank you. Ethan Halm? Ethan Halm on the phone?

Ann Hendrich? Ann Hendrich?

(Thomas Lee)? Is (Thomas Lee) on the phone?

(Jack Needleman)? (Jack Needleman)?

Marykay O'Neill?

Marykay O'Neill: Hi I'm the Chief Medical Officer for CIGNA in the Pacific Northwest. I did disclose on my form that I have a brother who's a Health Analyst for Group Health of the Pacific Northwest.

Ann Hammersmith: Okay, thank you. David Penson?

David Penson: Yes, David Penson Vanderbilt University Medical Center, nothing to disclose.

Ann Hammersmith: Okay, thank you. Doris Peter?

Doris Peter: Hi I'm Doris Peter. I disclosed on my form that in my work with Consumer Reports, we license information from some organizations that will probably submit measures like NCQA and other groups like that.

Ann Hammersmith: Okay, thank you.

Steve Phillips? Steve Phillips?

David Redfearn.?

David Redfearn: This is David Redfearn. I'm with WellPoint. I have nothing to disclose.

Ann Hammersmith: Okay, thank you. Jeffrey Rich? Jeffrey Rich?

Jeffrey Rich: Sorry, I was on mute.

Ann Hammersmith: Okay.

Jeffrey Rich: Sorry, I was on mute.

Ann Hammersmith: Okay.

Jeffrey Rich: Yes, this is Jeff Rich - sorry. I'm with the Mid-Atlantic Cardiothoracic Surgeons. I think I reported that I'm a former government employee. I ran the Medicare Fee for Service...

Ann Hammersmith: Okay.

Jeffrey Rich: ...System for the Bush Administration. But I really have no conflicts here.

Ann Hammersmith: Okay, thank you. William Rich?

William Rich: No conflicts to disclose.

Ann Hammersmith: Okay. And who are you with please?

William Rich: I am with the American Academy of Ophthalmology, Northern Virginia Ophthalmology Association.

Ann Hammersmith: Okay, thank you. Tom Rosenthal.

Tom Rosenthal: Tom Rosenthal, UCLA School of Medicine. I have no disclosures.

Ann Hammersmith: Okay. Barbara Rudolph?

Barbara Rudolph: Yes, the Leapfrog Group. We have three measures that were resource use measures, severity adjusted length of stay. But I have no other disclosures.

Ann Hammersmith: Okay, thank you. Joseph Stephansky?

Joseph Stephansky: Yes, I'm with the Michigan Health and Hospital Association. I have nothing to disclose.

Ann Hammersmith: Okay, thank you.

(James Weinstein)? (James Weinstein)?

Dolores Yanagihara?

Dolores Yanagihara: Hi, I'm with the Integrated Healthcare Association in California. We actually have Total Cost of Care measure that we're using. We worked with Thomson Reuters to develop it. We also - I also negotiate contracts with NCQA. And work on resource use development for our program and also with Thomson Reuters.

Ann Hammersmith: Okay, thank you. Has anyone joined the call since we started the disclosures who hasn't disclosed?

Ann Hendrich: Anne...

Ethan Halm: Yes, this is Ethan Halm, from University of Texas Southwestern. I have nothing to disclose.

Ann Hammersmith: Okay, thank you, anybody else?

Ann Hendrich: Ann Hendrich, Ascension Health, nothing to disclose.

Ann Hammersmith: Okay, anyone else? No one else? Okay, do any of you have any questions of each other or anything that you would like to discuss regarding the disclosures?

Okay, thank you very much. Have a good meeting.

Ashlie Wilbon: Okay everyone, we're going to, just give a quick review, as fast about review of the goals of this call and what we hope to ((inaudible)).

We realize that, you know, this process is new to a lot of you. Our phase one was very oriented around doing some thinking - broader thinking about resource use.

And moving into phase two it is a very specific process oriented thinking. So we realize that transition is going to be a challenge. And, you know, trying to follow the process and the evaluation criteria. And making sure that we are coming out with some ratings and rational for the measures that we evaluate.

So and through this process we are really looking forward to learning from any of the questions and things that come up in term of process. As well as ways that we can carry forward to future meetings to help things go a little bit smoother.

So we're not expecting at the end of this call that we come away with a recommendation from the steering committee on this measure. We realize that this is going to be the first time. And there's going to be some discussions that need to go on, in order to get comfortable with the process and the measures themselves.

So we'll have until the end of June, until after the steering committee in person meeting to actually make recommendations. We are already planning on scheduling a subsequent

conference call to discuss the Ingenix measure. And potentially continue discussion of this measure. Depending on how far we get on this call.

So really take this opportunity to ask questions of the developer. Staff is here to answer any process questions. But really just get comfortable with applying the criteria to the measures as you go through the process. (Sally), would you like to add anything to that?

(Sally): No. So as Ashlie said, we do have the developers on the call. So they're going to present on the measure and have the opportunity to answer any questions.

And we anticipate subsequent calls. We'll see how far we get through the four evaluation criteria today. And if we need to have a follow up call or add it to the in person meeting in June, we will do so.

Ashlie Wilbon: So we just wanted to review. Hopefully everyone is pretty familiar with the criteria at this point. For those of you that were able to complete some of the evaluation of the measures up to this point. But again, just to revisit.

There are four criteria of the measures -- Importance to Measure and Report; Scientific Acceptability of Measure Properties; Usability and Feasibility. We do expect that most people will find that the measures are important. It's a very focused - resource use is a pretty focused topic

And we realize that, that most of them, you guys will deem important. Anticipate that most of them will lie within the Scientific Acceptability of Measure Properties and understanding the measure construction and so forth. So just keeping in mind those criteria as you move forward.

So as I mentioned before we are looking for a systematic criteria based evaluation of the measures and thought process. We're going to start with the importance criteria and sequentially address the remaining criteria.

So once you've deemed the measure to be important, you will move onto Scientific Acceptability. And discuss those sub-criteria on to Usability and Feasibility.

We're really looking at the sub-criteria to delineate how the major overall criteria were met. So those sub-criteria are giving you clues as to how well the overall criteria has been met. So how do you know that the measure is important or scientifically acceptable etcetera?

And you'll notice also for those of you that were able to get into the SurveyMonkey and do some evaluation that the sub-criteria are all a degree of, rather than all or nothing. So it's not that - it's either yes or no, they met the criteria. But you are able to give a high, medium low rating of how well they demonstrated that based on the information that was submitted.

The Importance and Scientific Acceptability criteria are Yes or No. The overall criteria are Yes or No, but the sub-criteria can be rated High, Medium, Low. The Usability and Feasibility criteria, the sub-criteria, and the overall criteria are rated High, Medium and Low.

So and we'll get into more detail as we go on. But just so you have an idea that, you know, it's more of a bigger broader picture as opposed to a black and white picture of whether or not they did or did not meet the criteria.

So that said, I'm going to go ahead and hand it over to HealthPartners to begin. I'm going to bring up your slides in just a second. Is Sue Knudson there?

Sue Knudson: I am, Ashlie, thank you.

Ashlie Wilbon: Hi. Okay, give me one second to bring your slides up.

Sue Knudson: Okay.

Ashlie Wilbon: And can everyone maybe give it a second for your screens to catch up in case. But everyone should be seeing the HealthPartners orange slide on their screen and...

Male: Yes.

Ashlie Wilbon: Sue, just let me know when you want me to move a slide for you. Before you start though, does anyone have any questions from the steering committee about anything that (Sally) and I discussed before we move in to the HealthPartners' presentation? Okay great. Sue, take it away.

Sue Knudson: Okay great, thank you. Good afternoon everyone and thank you very much. To both (Sally) as well as Ashlie, the NQF Staff, as well as all of you on the steering committee for making this call for Non-Condition Specific Cost in resource use measures.

We here at HealthPartners have over a decade of experience using this type of measurement to improve affordability in our community. And it is with our pleasure that we offer it for endorsement by NQF for broad use.

So, on this note we want to start just by giving you a little bit of background on HealthPartners so you understand the context in which we've created the measurement technique and also have used it in operations of both our medical group and our health plan to target opportunity areas to facilitate improvement.

So, next slide please.

HealthPartners is a non-profit, consumer governed, integrated delivery system and financing system based here in Minnesota. We operate a health plan with over about a million members.

A care delivery system made up of our owned and operated HealthPartners Medical Group with nearly 700 primary care and specialty physicians. As well as Regions Hospital, which is a Level One Trauma Center and Urban Teaching Center in St. Paul.

In addition to our owned and operated delivery system. We partner with every physician group and hospital system in our network, which is not 100% but very close to all of the providers in our state and in the bordering states.

And we provide all of them with some level of measurement and reporting. It's our mission to improve the health of our members, patients and the community. I also might add to this background that while our market is highly competitive it's also collaborative.

We operate in a transparency enablement environment with a rich history of pooling data among payers. And most recently using some direct data submission techniques from care delivery electronic medical records system to provide quality performance measure information.

Our community is adopting standardized measures of patient experience to compliment these clinical measures. And those clinical measures, many of them begin with HealthPartners as the measurement developer.

We've put our algorithms and measurements back out for adoption of community standard. Culminating in them really becoming community property. So it is with that background that we offer our approach to cost and resource use measurement for more broad adoption.

Next slide please.

This is what I'll cover today. And it sounds like it is in synch with your logistics. So what is the measurement? Why is it important? How does it work? And how is it used?

Also we have provided the micro Web site link for your reference.

Next slide please.

So what is it more specifically? The Total Cost of Care Measure is a population based and person centered; and primary care focused measurement system.

The approach quantifies the effectiveness of managing a population of patients by examining 100% of the care provided to them.

It is illness burden adjusted, which is key to benchmarking. And also key to understanding relative performance. And in our submission we've provided output of both reliability as well as validity testing, which I can field questions on during the Q&A with the assistance of a few of my technical staff that are here with me.

In addition to that and maybe more importantly we've also, what we've done are what we call, "reality tested" it. That is, we've currently got about (70)% of our Health Plan Commercial Paid Claims on contractual arrangements that are based on this methodology. And we share savings with providers based on improvement of affordability. And we use this as a basis to dialog with providers in select areas of opportunity to improve the results.

It's also probably worth noting that we created this measure in partnership not only with our own medical group, but with others in our community and we've used the collective input of these groups to hone the reporting capabilities to be practical and usable in operations.

So finally we see this measurement technique as a technique to balance the triple aim. As I mentioned in the intro, quality and experience measures are somewhat abundant but affordability measures, at least a standardized approach is not.

So to some degree as well, we know that every health plan is already measuring illness burden adjusted PMPMs so transitioning it to an index performance with a resource use compliment, to us seems like a natural evolution.

Next slide please.

So why is this important? Briefly, as Ashlie had mentioned in the tee-up, we see that issues of affordability as paramount as supported by several of the citations that we included in our application.

I won't hit on them specifically, but we see affordability increasingly discussed, yet even in spite of that there's few publicly available cost measurement approaches that exist.

And so we also see this measurement approach as an opportunity to facilitate innovations both in payment reform, which I just mentioned our approach to shared savings agreements. As well as in new innovations in operational models such as the Carnival of Care Organizations and the advancement of those.

So we also see a very unique feature of this as a well rounded measurement approach to recognized population health efforts in prevention. As well as caring for those with chronic and acute needs.

So it really wraps in the total population ((inaudible)). And to that end it also importantly addresses the right care at the right time. And appropriately married with analyzing quality and experience measures. It can bear out issues of overuse as well.

Next slide please.

So now on to how it works. As mentioned earlier, at its core this is simply an illness burden adjusted PMPM reported as an index. It does rely on attribution, which in our mind is ripe for standardization. But in the meantime can be customized and is flexible.

We do use the majority of office visits to attribute. And we find these results from this attribution method synch up quite well with our medical groups view of their panel.

And assigning meaningful and accurate peer group cohorts such as a standard definition of primary care or how you're defining primary care is also a key component.

And then lastly, robust risk adjustment to capture the disease prevalence of a commercial population is also key.

So what's unique about this measure is a couple of things that I had already pointed out. It is population based. It is not a roll up of chronic disease nor is it a roll up of episodic care. So it does value the attention to disease prevention of a population. And appropriate resources provided to those with chronic disease and acute illness.

At the same time it does assure accountability for the volume of resources which sometimes we hear refer to as churn.

Then total cost of care lastly is a function of price and resource use. And when we talk about resource use we refer to the volume and the intensity of the services provided.

And as a part of this measure, we do have a unique approach for sub-setting resource use. Our approach puts 100% of care on a linear scale and values are relative and additive across sites of service.

We think that this approach is superior to re-pricing to a standard fee schedule among the sectors of care. Because service choices are put on an even playing field from a resource use point of view.

So I include a couple of examples here where - so what I'm saying here is, we can have an ENM code in the professional realm relative to a chest pain admission for example. And off this slide, but another example that makes a lot of sense to me sometimes is if I get an outpatient scan in the hospital, it is generally a higher cost than if I get it in a free standing center. But it is essentially the same service with a price or cost differential. And its - we're not viewing that as a resource use differential. And this methodology accurately portrays that difference.

And then lastly we recommend displaying it as an index for benchmarking.

Next slide please.

So here is basically how it works. The total cost index is a measure of the Primary care providers risk adjusted cost effectiveness at managing a population. And it's essentially their risk adjusted PMPM divided by that of the peer group.

And likewise the resource use Index is a risk adjusted measure of the frequency and intensity of all the healthcare services utilized in managing those patients. And also likewise it's the risk adjusted resource use PMPM divided by that average of the peer group.

Next slide please.

So now on to how it's used. In practice we segment the health plan information by primary care group and provide index performance reports to all of our providers. We provide aggregate information and a suite of metrics broken down by condition and by what we call, "service category". The next couple of slides after this one, I will go through a couple of examples of that information.

But other practical uses include the use in payment methodologies, like shared savings methodologies that I had referred to earlier. Also, benefit design and consumer transparency displays. And when we use cost of care measures in consumer transparency displays, we do couple that with quality and patient experience results.

So other potential applications though could be rolling it and looking nationally among regions in the country. Also, looking at hospital referral regions or hospital services areas for comparisons. However I note these as potential areas, but as it relates to our application, we have not yet tested these. And we also need to do a little work to create factors as was recently called out, in the JAMA article around differences in propensity to diagnose.

Next slide please.

So this is a bit of a detailed view, but let me walk you through. We wanted to give you a flavor of how we use this and how it's taken in action. So this is just sample data for illustration for you.

So the top box in the view contains some high level performance information over three years showing a number of our members attributed to this mock up provider, as well as the metro average. So you see the membership over 2008, 2009, and 2010.

Their corresponding Risk Score or Adjusted Clinical Grouping Score over the same time period. And you see that roughly about the same. Their total cost index which is below 1, whereas the metropolitan average is at 1. And this provider is performing 3 points better than the average.

And then our resource use Index. And so for 2010 you see that TCI at 0.97. But consistently across all three years the resource use a little higher than the metro average with 2010 being at 1.05.

So the next natural question might be, "Even though I'm performing well on the whole, are there still some areas where I could be improving?" And we provide drill down information using the same indices so its just really ((inaudible)) measurement to help guide that kind of discussion and improvement journey.

So you can see in the next level is where we start to break out detail by service category. And I'll focus on inpatient just for brevity in the discussion. But inpatient in 2010 you can see we have TCI of 1.04 and then if you look at the yellow box you see that inpatient resource use at 1.10.

And so that is quite a significant increase. So that is one thing to look as to what might have changed over those time periods. But then also it's high in the current time period so what might be driving there.

And that's where we have listed in the pink box on the bottom. Some high cost utilization measures that significantly drive these overall categories.

So from here you can see that this provider in this example has an admission count index at 1.02 so that could be one driver. Additionally inpatient surgeries are indexed at 1.08 which is meaning that they have more admissions and more surgeries.

I also might note that even though on the whole their outpatient TCI in the middle box looks good on the whole. ER Service and High Tech Radiology Services are a part of that. And you can see in this example that High Tech Radiology Services delivered in the ER as well as outside of the ER are also potential drivers and areas for opportunity.

So with that - next slide please. This slide shows now how the data can further be summarized by chronic condition. So any risk adjustment technique we understand allows for roll up of chronic conditions.

And again this method - we included this slide to show how the information and the method is drillable to chronic conditions. Yet flexible enough to measure the entire population.

So in this example you can see by these chronic condition and we have all other conditions listed on the bottom of the list. That you can see the members Associated, the total cost and the index associated with those conditions and the corresponding resource use Index.

And you can see that resource use when TCI is higher than 1.0 or higher than average, that resource use is a driver of that. And then also you can see that resource use continues to contribute, even when the overall TCI is lower.

Male: We're one slide behind aren't we?

Sue Knudson: Slide 10. No this is the right slide. At least - one back, thank you. So interpretation of the condition specific information I just say...

Male: There we are.

Sue Knudson: ...in conjunction with the quality results and results at a population level, can be easily portrayed in scatter plots. So we find for those people in using a population base methodology. Just looking at the condition specific information might not tell the whole picture.

And so within this context you have the opportunity to understand if people are enjoying better health outcomes with higher quality results and therefore able to address secondary health issues. And that might be leading to sometimes greater than average per person resource use.

And so this - we feel this nesting capability is essential to honing in where actual opportunities are and for it to be actionable at the provider level. And so we feel strongly that its only within the context of this population measure that you can do that complete and accurate assessment of condition based care.

Next slide please.

So this is the last slide I had. And not to belabor it, but just to show you our own Partners medical group has used this over the last several years to drive improvement. And as many of you may be aware they are both award winning with their care model design or care model process. As well as cost effective with their result.

And this list simply shows some of the areas of focus that have been directionally informed by using this type of measurement approach.

And with that I just want to thank you for the opportunity to provide this overview and I'd be happy to take any questions. Thanks.

Bruce Steinwald: This is Bruce Steinwald. I have two questions for you. You submitted the two measures simultaneously the TCI and the RUI. And it is probably the RUI that is most germane to what we've been doing. We're been looking at resource use, not cost measures.

Is it right though to say, that the only difference between the two is that the total cost index factors in price variations across providers. Whereas the resource use doesn't.

Sue Knudson: Correct.

Bruce Steinwald: Okay. Second question, is your resource use index relies on a total care relative resource values. And you've managed to construct those values so that they are additive across different sites of care. Using RVUs and using other measures developed by Medicare and then adding in pharmacy.

Can you tell us in 25 words or less, or more if needed, how you did that?

Sue Knudson: Right, so I will give it a shot. And I also - if you still have questions I can rely on some of my technical team members who are here as well. We do rely on those sector specific...

Hello?

Bruce Steinwald: Hello?

Sue Knudson: Yes.

Did someone just put us on hold maybe?

Thomas - Tom are you there, operator?

Anyone else there? Bruce are you there?

Bruce Steinwald: I'm here.

Jeff Rich: I'm here, Jeff Rich.

Ashlie Wilbon: Hi, I'm not sure what just happened. Tom? Operator, are you there?

Male: Wait, I think we lost our speaker.

Sue Knudson: No I'm here can you hear me.

Ashlie Wilbon: I'm here. This is Ashlie from NQF, I'm not sure what just happened. But...

Bruce Steinwald: Somebody went on mute. It happened last week too.

Female: Or on hold. Yes.

Bruce Steinwald: Hold, yes.

Ashlie Wilbon: Sorry about that. Please don't put us on hold.

Sue Knudson: Ashlie, can you hear me?

Ashlie Wilbon: Yes, I can hear you.

Sue Knudson: Okay, very good. So in answer to the question about how we construct HPRRV. We do rely initially on the sector specific relative value units.

From there we look at the bill charges across those different sectors of care to build relativity. And then we do a final step to look at the paid amounts within sectors to make sure that it is adjusted appropriately.

And then we have some final quality checks for threshold and outputs. Like code level specific areas, that might have counter intuitive results you might say and do some final checking there. So that is the very high level and we can get into more detail if anyone is interested.

Bruce Steinwald: Well, but the result of it is that you have all types of services and sites on the same scale. That is that you've got a follow up office visit performed in the physician's office on the same scale as an inpatient a service on the same scale as an outpatient diagnostic test. Is that right?

Sue Knudson: Yes. That's right.

And I would also add. I think we may have made note. But we did - we are attempting to patent this, but not for commercial licensing. But only to - we are interested in sharing it for community property.

Bruce Steinwald: Okay. Other steering committee members do you have questions?

David Redfearn: This is David Redfearn. The one question I think is interesting is that you use billed charges when you use relativity. Does that bring in some distortions based on the sort of the way people - the way bill charges don't actually reflect the cost of the service?

Sue Knudson: Yes, that's right. But what we found was that two things, is that also the paid amounts even among providers sometimes, and where things are loaded also is not as rational. And so we found that using the bill charges in that first step. And then calibrating among the sectors of care was the best way to account for those confounders.

Marykay O'Neill: This is Marykay O'Neill on that same topic. This got me wondering if this varies by market with how much distortion those charges will introduce into the analysis.

Ashlie Wilbon: Hi, this is Ashlie. Sorry to interrupt. I just want to bring us back before we kind of skip around a little bit. I have a feeling that a lot of these questions are going to come up in our discussion of Scientific Acceptability. And in the spirit of trying to follow the process we're going to...

Female: Okay.

Ashlie Wilbon: ...bring you guys back to importance. And have Doris and Bruce kind of start the discussion from there. So please hold your thoughts. Write them down if you need to.

I realize that Sue just presented and there's a lot of stuff right on the tips of your tongue but we just want to make sure that we stay within the process as much as possible.

So I'm actually going to bring up the document that we sent out to you guys. As a guide for how to evaluate the measures based on the criteria. So I'm going to just put that up on the screen for the screenshare. And then hand it back over to Bruce and Doris to hopefully guide you guys

through the discussion of importance and I'm sure we will make it into Scientific Acceptability before we ((inaudible)). So I'll hand it over to you guys.

Bruce Steinwald: Alright so we save questions for when we're discussing the criteria. All right. Then let us start with importance. Would any steering committee members like to raise any issues related to importance or ask any questions of the developer?

Doris Lotz: Yes, Bruce, this is Doris. I wanted - before we take off and kind of go line by line and go 1-A, 1-B, etcetera and try to capture some sort of straw vote and discussion. Since - I got hung up on your comment about total cost index and since we're not considering cost but rather resources, or are we considering cost? Are we considering two measures here or are we considering only one? How - that's a question for the NQF staff. How do we proceed with this? Are we doing this separately or as a group? Are we taking the cost out?

(Sally): This is (Sally). That's a great question Doris. The way that we've presented it - the costing method approach was left open to the measure developers. So a measure developer could potentially bring through a measure that applies actual or allowable charges, some kind of - versus the standard price or just straight utilization.

To the extent that the steering committee thinks that these are two measures, I think we need to talk that through with the developers and see if they really consider them two measures or is it different parts - different points in estimating the measure. I think that's a great question, whether it's one or two measures. But both potentially could come under this umbrella -- both approaches.

Doris Lotz: Again this is Doris. I - to me they seem like two separate measures.

(Sally): Okay.

Doris Lotz: Certainly proceed independently. They could be presented independently.

Bruce Steinwald: This is Bruce. I mean - to be consistent with all that we have done before, it seems to me that it's the resource use index that we should be looking at.

Jeff Rich: Yes. This is Jeff Rich. I agree completely.

William Rich: Yes, (Bill) Rich, agree.

Barbara Rudolph: This is (Barb) Rudolph. I actually disagree and see that the total cost of care is sort of where you get to and that the resource use is something that you can drill down to as a way to understand the cost of care. But I'm not the measure developer, but the measure developer did give it the title, "Total Cost of Care and Research Use Population Base PMPM Index," which did not necessarily imply that they were to be separated.

Bruce Steinwald: Well then, let's hear from the measure developer.

Marykay O'Neill: Thank you. Yes, we agree they can certainly be used independently, but our intent and our strong feeling is they are better used in partnership with one another.

Male: Okay.

Female: (This is) Marykay O'Neill, I'm ((inaudible)). This is an area of considerable importance to me because I think it's something that if the cost is reported, it has to be done very carefully to actually convey real information. On the other hand, there are aspects of cost that I think are important in information about how systems are working.

For example, I'm not sure that a pure utilization measures something like use of radiologic studies in different locations, the various economic impact of that could be captured without some cost component, you know, unless you're counting facility fee lines or something like that. So, I guess I'm arguing both sides of the coin here but I am very troubled about using charges. Based on experience in my market, looking at that kind of data, it would be a very hard thing to ((inaudible)).

William Rich: This is (Bill) Rich. And I agree with ((inaudible)). You know, I think we're going to have to add some ((inaudible)) here before we can ((inaudible)). I didn't quite understand your methodology when you said ((inaudible)) charges and then somehow you adjusted for those charges - it really wasn't clear to us, at least to me, in your presentation what you meant by that. Otherwise, you ((inaudible)). How can this be used in different markets ((inaudible)) different populations?

(Sally): William, this is (Sally). You're cutting in and out about every second or third word unfortunately. I don't know if you're having reception challenges.

William Rich: I can hear you fine, but -- is my voice still breaking up?

Male: Yes, it is.

William Rich: I'll hang up and get on another line.

(Sally): Okay.

Bruce Steinwald: As I understand it, the use of charges though would apply to both measures, right? The charges were used in order to create a single scale for every kind of utilization and that would apply to both measures?

Susan Knudson: This is (Sue). In response to that, we use the actual paid, so the allowable inclusive of member liability for the total cost measure. And on the resource use measure, perhaps I'll need to repeat this when the other gentleman is able to rejoin us...

William Rich: I can actually hear you fine, go ahead.

Susan Knudson: Oh, okay very good. We use the billed only to understand the relativity among the services. But we use an adjustment factor for paid-to-billed so that we get that relativity adjusted appropriately among sectors of care. For example, in-patient to out-patient to professional services.

Bruce Steinwald: You use an adjustment that makes it - it sounds like you said it's been based on actual payments as opposed to charges.

Susan Knudson: Well, we want the paid adjustment to understand the overall spend, what proportion of it is in which sector of care and that's where we do that across sector of care adjustment versus among the services, if you think of it that way.

Doris Lotz: So you're applying a different - you're applying some kind of adjustment factor because discounts are different in different types of care, in-patient, out-patient, professional?

Susan Knudson: I'm going to actually ask one of my colleagues, ((inaudible))...

Doris Lotz: Okay.

Susan Knudson: to comment to see if he could say it maybe a little more clearly.

Male: Thanks Susan. I'll do my best. So what we end up doing to do this, we actually use the billed amount to apply it to the lines or to the claims so that's the most standard piece of information we have coming from the data. So we apply it to the lines using the billed amount. But when we go across the different components -- like in-patient, out-patient, professional -- we know that there's a different discount rate for each of those components, so we want to put it back in line to the real world.

So, the adjustment factor we do adjusts all those bills back down to the paid, so what you - in essence what your paid ratios would look like, your billed ratios would not. So we want to make sure we reflect what's really going on in the real world. So the billed amount is only used to apply to the lines but the paid adjustment factor gets its back to the real world, across components.

Doris Lotz: Do you adjust the same adjustment factor for different types of service on some kind of community average or is it specific to payer?

Susan Knudson: Well, we are one payer, so it's specific...

Doris Lotz: So you're - it's all internal for you?

Male: Correct.

Doris Lotz: Okay.

Susan Knudson: That's correct.

Doris Lotz: That makes it easier.

Susan Knudson: But we've also - we have - we developed this methodology by use of a national database and so it's - we have waste calibrated nationally and regionally.

Doris Lotz: Calibrate on national and regional average discounts?

Susan Knudson: Yes. So it's not HealthPartners specific. That's available.

Doris Lotz: Okay. I'm - I guess...

Male: Is the...

Doris Lotz: Okay, I'm done.

Male: Is the attributable population HMO patients or PPO patients or everybody?

Susan Knudson: For the attributed membership, we're looking at -- now you're off the HPRB question and you're on to attribution? I just want to clarify?

Male: No, I'm sorry. I've probably skipped ahead.

Susan Knudson: Oh. But, for attribution, we're attributing within our population as a health plan. And, what we're saying is that can be scalable to different units of analysis.

Male: So your health plan has HMO and PPO patients?

Susan Knudson: Yes and it's essentially the same.

Male: So what's the rule set for attributing it then to primary care physicians?

Susan Knudson: We look at the practicing specialty of the physician and we look at claims history for evaluation of management codes and we attribute patients to the clinic with the majority of visits.

Male: Okay.

Bruce Steinwald: We still need to resolve the question that Doris raised. Is, are we - it sounds like HealthPartners you'd like us to evaluate these two measures as a, basically a combined measure. You said that you think that they work together and you'd be uncomfortable with us evaluating just the resource use measure, is that true?

Susan Knudson: Well certainly you could evaluate the resource use measure. We think to the comments made by a few of the previous individuals in the example that I gave in the presentation that that is really where the more action ability of the measure comes out in play is that if you just had a resource use measure, like in the scan example that I gave in the presentation and one of the steering committee members had also referenced -- because we value that from a resource point of view the same, we're really not giving credit, if you will, to the provider who's from a stewardship point of view using a more cost effective venue for the service.

And so, that's where we think the effectiveness of using ((inaudible)) measures in combination and its ability to roll down into details and roll up into the population is really where the strength of the tool and that approach is.

Bruce Steinwald: Well you've given an example where you have a procedure that can be performed in a different setting and it's likely that the cost will be higher in one setting than another, but what about just price differences across different provider groups for, you know, hospital and other services? Don't those also influence the cost measure separate from the resource measure?

Susan Knudson: Yes, that's true. And in fact, operationally, you know, we have a price index as well, but we knew the call was not for the price index and so, but that's the way, you know, this market has advanced to that level of discussion and so yes, that is very much in play.

Bruce Steinwald: Okay.

Jeff Rich: So, this is Jeff Rich. So, I'm a cardiac surgeon and I'm operating in every one of your localities and are you looking at the resource utilization that I use versus everybody else uses within your network regardless of geographic area? And then, if you are, how do you adjust for the geographic variation and the competitiveness of the marketplace for cost, because you are discounting services in some competitive markets where you have a strong competitor and in others where you are a monopoly, you probably aren't discounting them very much.

So, the difference - are you looking at providers based on just a bucket of resource utilization? Do I do as many chest x-rays as the next guy does, you know, 500 miles away from me in Minnesota? Or, am I over-utilizing that resource? And then, how do you control for that resource based on the cost and the market competitiveness?

Susan Knudson: So, I'll try to answer all of those, but if I miss anything, just let me know. And, a couple of things. First of all, when we set up peer groups and do profiling for benefit design or product development, we do it within meaningful competition areas.

So, for example, the displays that I provided were based on our metropolitan area which is the twin cities of Minneapolis and St. Paul and the surrounding suburbs. We would analyze the northern Minnesota Duluth area, for example, as a separate sort of peer group for consumers in that market area.

But again, the flexibility of the tool helps one understand. Like a purchaser, if you've got - if you've got an employee base both in the twin cities as well as someone - an employee base in Duluth, it helps you also understand the relative differentials between those markets. So depending on the application and the user, it's got that scalability and flexibility. Did that hit on your - all your...

Jeff Rich: Yes. So I mean, I like the idea of the purity of a resource utilization measure bucket. So I like the idea that you would compare me to somebody else in northern Minnesota for delivery of the same set of services based on resource utilization and that would be number of CVCs and chest x-rays and tests that I order.

And then, my question to you is, as a paired measure then, you know, do you feel like you have a fair and balanced way of looking at the cost of those resources based on geography and cultural differences? Because I thought when I was at CMS, the differences between - the cultural differences between coronary bypass grafting in Minnesota and southern Florida were enormous because the number of consultants in southern Florida was three times the number of consultants in Minnesota on a per-beneficiary basis.

And so, we sort of looked and scratched our heads and realized that there are true cultural differences in the way people deliver care in their community, and their expectations, both by patients and other providers. So, how do you account for that?

Susan Knudson: Well, I think in the example you're providing, that would bear itself out in differences in resource use, if there is different practice patterns, that would come out in a resource use index. And then locally, if you've got the same providers -- so just to take that example a step further because I think I may have heard this in your initial question -- is, even in this market, we've got multiple hospitals and we might have practices that do admitting to multiple hospitals, for example, but they might have different price points.

And so, there again the resource use consumption might be somewhat consistent for any physician, you know, with a patient whose - or a part of their panel that has surgeries, but what might be different among those different hospitals is the price they charge. And again, that would come out in this pairing of measures as the total cost index might be a bit higher but the resource use might be lower.

Jeff Rich: Thank you.

(Sally): So, this is (Sally). I just want to interject really quick a process point. For this particular project, and it sounds like the conversation is going this way, we are accepting individual measures but we are not accepting paired measures for evaluation. The project and the - even the databases - the evaluation criteria are not necessarily setup for that.

I don't want to stop the conversation, but I just - we need to make a decision with a very astute question about whether this is one measure, whether these are two measures that the developer would like paired, which means it needs to - we need to rethink about how to manage that and when to manage that in an evaluation process.

Susan Knudson: (Sally) it's Susan. I just want to clarify, the measures are not dependent on one another.

(Sally): Okay.

Susan Knudson: But they - so they can be used independently. There is added value when they're used together.

(Sally): Thank you. That's very helpful.

Bruce Steinwald: But (Sally) you just said that we're not accepting paired measures, so our choice here is one measure or two measures. If there are two measures, do we have to evaluate them separately?

Susan Knudson: Yes. They would have to be evaluated separately.

Female: Even if they were paired...

Susan Knudson: Yes. In fact, even if they were paired, they would be evaluated separately and then, you know, how they work together. So...

Bruce Steinwald: Well in that case, I go back to where I started, which is, I think that we ought to be evaluating the resource use index and if we decide we want to also evaluate the total cost index then we can do that. But, if we have to go through the process separately, I would prefer to go through the resource use one first at least. Does any other member of the steering committee have a view?

Doris Lotz: Yes, this is Doris. I would agree and I was just waiting for a pause in the conversation to say in the hour that remains, or for me, a little over a half an hour before I have to jump off, I think that it would be good both to apply the process that folks learned about a week or so ago to just the resource measure right now.

We can learn certainly about this resource measure while we've got the developer on the phone with us, but we can also kind of apply this relatively new process for folks over this conversation and see how that works. So if we want to modify how we approach all of the other measures going forward, we can learn something in this conversation and make it a more efficient and effective process after this conversation.

So, I'd like to strongly urge us to just charge forward with the resource use, you know, going one A, one B, and get as far along as we can and then see what we've learned at the end of the hour as far as how to structure the next couple of phone calls and ultimately our meeting in June.

Jeff Rich: Yes. I would agree with that. I think that resource use measures are foundational here and that's what we're all about. And then we can move on from there. But, I think we'll have enough work just looking at resource use measures.

Bruce Steinwald: Yes. So, why don't then, without objection, why don't we go back to importance and start looking at the specifics of the sub-criteria? (Sally), is that what you would like us to do, to go sub-criterion, one at a time?

(Sally): That's right. So, in this instance, because the non-condition measures have not gone to a technical advisory panel first, they go straight to the steering committee, we need the steering committee to focus on the sub-criteria to start with. And, you know, keeping in mind what the goal of the over-arching criteria.

So, for importance, is this area in which the measure's targeting something that's important to measure and henceforth while the resources that it would take to measure. So you'll see the sub-criteria, one A, one B, one C on the left-hand side on the webinar, or if you're not have access to the webinar, it's the evaluating resource use measures table that was sent out to all of you.

Bruce Steinwald: All right then, I would like to propose that criterion one A is met because I think it's, the order -- I don't know about the national health gold priority -- but either one, I think is met by this kind of measure - this measure.

Male: Agree.

Female: Agree.

Male: Okay.

Male: Agree.

Male: Agree.

Female: Agree.

Doris Lotz: This is Doris. I'll agree with that with a comment. I think it's a teaser. You know, I think it's hard to do anything with this measure without substantive follow up. I think it's, you know, it's your ticket to the dance. But it doesn't tell you much beyond that. So, I agree that it does have measure, but I'm curious to see how our conversation unfolds from here on out because as an isolated measure it doesn't tell you much.

Barbara Rudolph: This is (Barb) Rudolph. I think it does tell, certainly the providers, areas where there is perhaps overuse or under use. I think it tells a lot. And, I disagree on that assessment, but I think it meets this first criteria given the fact that there is grave concern about the 30% of overuse and waste in the marketplace. So...

Bruce Steinwald: Yes. I'd certainly agree with - I mean, you know, we're holding off on the other criteria, so we're just looking at importance and I would agree, if you're at 1.15 and the other criteria have to be met as well, that's telling you something important.

Doris Lotz: I think it gives you a kind of detail that allows a clinician or clinical entity to target their evaluation in that very specific area. So, I think it's quite actionable on that level. I guess to do a - to see how it works on a larger rolled up level or community basis, it is not as clear.

(Sally): Okay. So, Bruce you want to move it to one B and one C?

Bruce Steinwald: Yes.

(Sally): I think some of what we're hearing is linked to feasibility which is fantastic in usability but if we can try and think about the topical area that is being addressed by this measure and whether it's - examining an area where there seems to be problems and an opportunity for improvement. Again, this is not about whether or not the measure itself can demonstrate that but whether or not it's examining an area in which we think there are problems.

Bruce Steinwald: Having said that, I'd propose that all of the sub-criteria, I believe are met. Is there any disagreement?

Female: No.

Female: No.

Male: No.

Bruce Steinwald: Then why don't we go on to number two.

(Sally): So just quickly - this is a process check. The rationale that all the criteria and importance have been met is because the measure's targeting area where there's known to be variation, relevant service categories, and that the objective of measuring resources have been clearly described. Does that...

Bruce Steinwald: I believe that the supporting information that the developer provided helps to demonstrate that. I mean we all know that ourselves.

Female: Right.

Bruce Steinwald: But they also documented that, I think, satisfactorily.

(Sally): Fantastic. Thank you.

Now moving on to where there's going to be a lot of discussion and we'll see how far we can get today is the scientific acceptability. And you'll note that there's a lot more sub-criteria to work through and also a lot more of the items in the specifications - that form that you had where in blue text is what the developer submitted, is being mapped back so you can see there's a lot of information to walk through and I'm sure a lot of questions to be addressed.

Barbara Rudolph: A question on this - sort of a content - I found the manual or the technical manual to be very complete in these areas so I would hope that that's considered as well because there were references all the way through and - to match up with what was in the technical manual.

Bruce Steinwald: Are you referring to what was available through the partners' Web site?

Barbara Rudolph: Yes.

Bruce Steinwald: Oh, okay.

Male: Yes, I'm having trouble getting some of those things downloaded, some of the files are shown as being corrupted or just...

Male: Yes.

Male: ...it won't download.

Barbara Rudolph: But the Web site is really very clear. If you have a chance to go there, the link is...

Male: Well, I'm not quite so sure - the documents that I have downloaded and have read, I'm not sure the clarity - I'm still struggling with the clarity on this.

Susan Knudson: If there's a technical issue with some of the documents we'd be happy to email them directly to anyone, size permitting. But, it sounds like the Web site's working okay for some but understand if that might have issues with some as well. I would just offer that (Sally).

(Sally): Thank you Susan.

Bruce Steinwald: Two A, in and of itself, is pretty complicated, or at least it's pretty inclusive of a lot of things. Would anyone like to start a discussion about how well-defined and precisely specified the measure is?

All right...

Doris Lotz: I guess - I have - I think I have an understanding about the measure development within the context of HealthPartners and I understand from an earlier comment that they have checked it up at least in some regards against some regional and national numbers, but I think, you know, as somebody that has been a practicing clinician, I find that there's a lot of actionable data in this reporting at that level.

I am struggling a little bit to understand how it could be transplanted to other systems. And, without a fair amount of work -- I guess maybe that's the feasibility to this thing -- but just how much of the measure's definition and precision is based upon the internal systems that within HealthPartners as opposed to what might be available on other systems.

And then in terms of the definition of the measure, if we're looking at this as a population measure, I guess there's a category of data that I think is missing which is, if there is - if there are individuals described on a population basis to a care system or a region or something like that, there was no way of looking at what it implies in terms of appropriate research utilization that you have people that aren't seen at all or their whole encounter is purely a pharmacy claim.

It seems to me that they're excluded but yet from a population basis there actually can be a risky group. So the non-utilizers haven't been defined in the data set for inclusion in an evaluation and then the specificity of the data that resides within their system being easy to obtain external to HealthPartners.

Susan Knudson: May I comment on that?

Bruce Steinwald: Please.

Susan Knudson: You know, we have done it - conducted use of the measure denominated by users using attribution, but it could very easily be denominated with population of total eligible, even recognizing some of those eligibles may only have pharmacy claims, as you say, or are not using any services. And that - so that is flexible. Again, I think it's defining the application and use and it's got a lot of flexibility. And then customizing those areas to meet that.

Female: It was my understanding that a pharmacy was only excluded by Medicare CMS in their calculations but that you didn't exclude a case that only had a pharmacy. Is that correct?

(Chad): This is (Chad) on Susan's team. I want to comment on that. No, this is applied for a commercial population.

Female: Right. I was just looking at the - some of the methodology for the total cost, TCRR. In there it said that you received some of the waits from CMS, is that correct? Or (relative waits).

Susan Knudson: Right. That's for the waiting within the sector of care. So, for example, the DRG waits, we would have used the CMS waits.

Female: Right.

Susan Knudson: But the calibration...

Doris Lotz: But - but my understanding with attribution methodology is those people would not necessarily be attributed.

Susan Knudson: That's correct. Non-users would not be attributed but if, depending again on the application, if you wanted to broaden the use to a total population versus those just using the system, that could certainly be done.

Doris Lotz: Okay.

Susan Knudson: That's where we said that attribution and some of these things are pretty standard, but can also be flexible to change.

Bruce Steinwald: Another question about attribution. So, I believe the materials say that the patient is either, has selected a primary care physician and that physician then is attributed that patient's

utilization, or a statistical measure is used and it's the patients whose - the doctor who has the most primary care visits with the patient.

So the patient has to be a user of primary care services to be included as I understand it. And then the attribution is at the individual doctor level but the peer group appear to be large group comparisons. Have I got that right? So you're not actually comparing individual doctor's performance measures, you're comparing the groups that the doctor belongs to?

Susan Knudson: Yes. Two things there. First of all on the attribution, we are operating in what we refer to as an open access market. This market in the commercial realm is not accepting of patient's assignment models. And so we have used a retrospective attribution. That said, you can again flexibly use a prospective attribution model as well. And so, for example, in public programs in this state, there is some products where prospective attribution is in place and so that could work as well.

But, we have predominately in this commercial evaluation used the retrospective model. And we are attributing at the group level. We have to look at the individual physician for the practicing specialty to make sure that we have an accurate definition of primary care when we're doing that attribution, but that said, we do not profile at the individual physician level, if you will. We profile at the group level.

Bruce Steinwald: All right. So then, all of the physicians who are included have to be members of groups that are large enough to have at least 600 patients or something like that, right?

Susan Knudson: That's right. So this market is very much a group practice, orientated kind of market and so those are the thresholds that we have tested.

Bruce Steinwald: Would anyone like to either speak in favor or against stating that they've met criterion

2a1?

Barbara Rudolph: I would speak in favor of it. This is Barbara Rudolph.

Bruce Steinwald: Okay.

Paul Barnett: This is Paul Barnett. One thing I was curious about is the decision to exclude these very high claims from the data. And, is that - now is it - it's a \$100,000 claim or is that individuals with \$100,000?

Susan Knudson: So it's not an exclusion, it's a truncation at the individual level.

Paul Barnett: A truncation, yes. So you're excluding all expenses above \$100,000 per year, per person?

Susan Knudson: Per person, correct.

Paul Barnett: And so, how much is excluded?

Susan Knudson: Pardon me? How much would the....

Paul Barnett: How much of the total amount is excluded thereby and why \$100,000 and why not something else?

Susan Knudson: I think I would need to follow up for the precise number, but it's about five to eight percent, maybe.

Male: Dollars, yes.

Female: Would that apply to a pure resource use measurement or only when you're doing the total cost?

Susan Knudson: It would be both.

Female: Okay.

Susan Knudson: We apply the truncation consistently throughout.

Female: So any services that hit after they hit \$100,000 worth of services are not counted as additional services?

Susan Knudson: Well the claims are truncated, consistently all the way through. And so you have - you still have claims, we just truncate them using a factor that applies to all claims all the way through. So, all the relativity still stays in tact.

Female: So you'd still be counting all the procedures and services?

Susan Knudson: That's correct.

Female: Okay.

Paul Barnett: So they're in the resource measure but not in the cost measure. Is that what you're saying?

Susan Knudson: They're in both.

Female: But the impact is really only on the cost side.

Bruce Steinwald: Well, but as I understand it, the example you gave, the \$125,000 patient's resources are all scaled back, so, tell me if this is right or wrong, the patient had an MRI, he's actually considered as having .8 of an MRI?

Susan Knudson: That's correct. Both are scaled using the same truncation factor.

Female: So it seems like you may not really be able to separate the two then because if you're - if you need to actually figure out the total cost of care in dollars to apply the truncation, I mean, how would you do that without doing the total cost first?

Susan Knudson: Well, and that's a key that consistency throughout the methodology is well, because we're, you know, in order to risk adjust your resource use appropriately too, you need truncation there as well. And so, you really do need the capability of doing that method.

And to the extent that, as I had mentioned, even if we're not saying these two measures were just evaluating resource use, you need full claims data to do this as well. And so, by virtue of that, if you look at health plans or payers as really the ones producing the information, they'll have that number there to be able to truncate the resource use number as well.

Female: It would require that the plans or whoever is doing the calculation to have the actual payment amount on the claim.

Susan Knudson: Yes.

Paul Barnett: And so, why - what was the rationale about excluding these very high cost patients or that portion of their claims above \$100,000?

Susan Knudson: Well, we don't want to bring in the is(Sue) of accepting insurance risk and then also, we were going by the guidelines that were published and we referred to in the Society of Actuaries report. And it's a pretty consistent truncation limit that's applied through several methodologies. We're not unique in this.

Dolores Yanagihara: Yes, we actually applied - this is Dolores with IHA - we actually apply the same level into our total cost-of-care measure so it's consistent with that too.

Female: I'm not sure if it comes here or in the maybe the feasibility or usability section, but I'm grappling with the same question that, I think, with Marykay that asked in how would you apply the (TCRRV)s to other regions. So if this is a standardized measure to be used across the country, how would that, you know, those values apply. I mean, would you just use the values as-is or would they need to be adjusted and how would you do that adjustment? I'm really grappling with that.

Susan Knudson: So the table that we had provided as a part of the application is just simply at the code-level so when they potentially -- use can just if you will, to oversimplify -- kind of glue it on their file or merge it and match it up. And we created that. We have our own calibrated for our own use but we also had access to an external database based on 20 million lives nationally.

And so we created national and regional (HPRRV)s for broader use.

Female: So someone from another region could actually just take whatever regional and apply it to them and use it as-is?

Susan Knudson: Yes.

Female: From that national database? Okay, great. Thank you.

Female: I guess I'm a beating a dead horse here, but on the (TCRRVs), I mean, I don't now if there are - that if you look at different regions or different markets, if there are some markets for which there is - I mean, any of them can have an average that some of them would have huge confidence intervals or spread as compared to others.

And we did a study on hospitals internal to the state of Washington and the bill charges are just, you know, every institution had a different strategy about how to set their charge faster so that it was all over the map. And an average really wouldn't tell you much about any systems behavior.

Maybe it's not as germane as I think it is but I think it is a little bit problematic.

Male: Well, if I've got this right, didn't you use Medicare's values? And the key is to get every site and type of service on the same scale.

But the relatives of, you know, for different inpatient hospital services, don't they come from Medicare's MS-DRGs?

(Sally): That's right.

Male: Okay.

Female: Okay.

Male: Well...

(Gary): I think they're just filling, doing the gaps from other sources as I understand it.

Susan Knudson: Yes and then also doing that step to calibrate across those different sectors of care -- inpatient, outpatient, professional, et cetera.

Male: Right, cause they're not all on the same scale, right.

Female: Correct.

Male: And the pharmacy relative values, where would they come from?

(Gary): This is (Gary) again. It is actually the billed amount of the pharmacy. So we just create our, since there's weights available, we just create them using the relativities of the billed amount between the NBC Codes and again, we calibrate that back to the inpatient outpatient professional pharmacy.

So it's an average billed per pill.

Male: Yes, so that's kind of interesting because my understanding is that the billed amount for generics as the wedge between the billed and the actual cost of the payer is much bigger for generics because of the rebates. An interesting problem. Maybe not soluble because the rebates are all hidden from view.

Susan Knudson: Right.

Bruce Steinwald: This is Bruce. I guess also we should be looking at their risk adjustment methodology in this section as well. It's listed under the submission items and I know there are some members that had some concerns, not so much about Partners, but about these different methodologies

that are commercially available. Would anyone like to make a statement either way?

They're using the ACGs developed at Johns Hopkins and the latest version, I presume.

Susan Knudson: Yes, we're using the latest version, 9.0 I believe.

Bruce Steinwald: Yes, so that is an interesting question because we're just sort of - I mean we don't really go into here what goes into building that right? We're just sort of taking it off the shelf.

Susan Knudson: That's right. In our experience, we're looking for (robust risks) given comparative nature and use of this kind of measure. And so, we need one that we can rely on to cover the disease prevalence of the commercial population and to really be accurate in terms of guiding improvement activities when it comes to care models and practices.

Bruce Steinwald: And you relied heavily on a Society of Actuaries study that had sounded like they concluded that a number of these commercially available risk adjustment methodologies are satisfactory for most purposes?

Susan Knudson: That's correct. And just to share a little bit of history for our market, we pretty much have a market standard history using ACGs. I understand more recently there might be some different ones used.

So we have - even our local department of Human Services, for example, uses ACGs to risk adjust. And so that's one aspect that goes into our use of the ACGs but we did note that any of the risk adjustors associated with the Society of Actuaries study, you know, are comparable as it relates to their findings.

Any ones that aren't on that, then we suggest that they be tested.

David Redfearn: This is David Redfearn and I agree completely. I mean, there's a bunch of these kinds of things that are all subject, basically to the same kind of limitations in terms of how much of the future costs variation they can account for. But they're all roughly equivalent. I don't think there's a lot to choose from between them and the model they've chosen, I think, is okay.

Bruce Steinwald: This is Bruce. Should we discuss more in this category or can we go on to reliability testing?

(Sally): Yes so to A1, this sub-criteria is about whether or not as submitted, the measures defined and precisely specified enough. So it's as you noted, a lot of the components are not applicable for this measure as demonstrated in the submission but if you feel that you've had enough discussion about this sub-criteria, which is about the specifications themselves. The next one, you're absolutely right, would be reliability testing information and results.

Bruce Steinwald: So I think we've lost our screen.

Ashlie Wilbon: Yes, can you - thank you. We weren't sure if the server disconnected on us for no apparent reason, we're not really sure, we'll follow up with the webinar folks. But we think we reconnected it so the same side-by-side table should be back up. Are you still not seeing that?

Bruce Steinwald: I don't see it.

Bruce Steinwald: I had to log in again.

Ashlie Wilbon: Yes, you might have to refresh. I apologize. We don't even know what that was. We got some strange warning box saying that we were disconnected and being reconnected. So we

weren't even sure if it was showing up on your end.

Please refresh. We do still see now the names of all of you in so it looks like it's just a matter of the presentation itself. If you are not able to get back in, we were on page 2 of the side-by-side table, looking at Sub-criteria 2a1, which is about the precisely defined and specified specifications to allow for implementation.

And then reliability starts at the bottom of page 2 and goes over onto page 3.

Is the refresh working for folks?

Bruce Steinwald: Afraid not.

Male: No.

Female: You have to re((inaudible)) go back out totally.

Male: I did that and it still didn't work.

Male: Yes, I did too.

Ashlie Wilbon: Okay. Yes. We can try and contact with the webinar people but I would hate to lose this time. But you have the email that (I actually) sent last Thursday, I believe it was, the side-by-side table is located in that email.

Male: Second of the - I think it's Wednesday actually.

Ashlie Wilbon: Wednesday, thank you.

Male: Yes so the question was 2a1, you were - what were we...

Bruce Steinwald: Well I'm not sure that we actually need to come to a collective judgment or not because we don't have to.

Male: Right.

Bruce Steinwald: So I guess I was asking if members are satisfied to move on from that topic and talk about reliability.

Female: Yes I am.

Male: Okay.

Bruce Steinwald: Okay.

(Sally): So reliability sub-criteria 2a2 and 2b1. Starting with 2a2.

Male: You know, there was one thing on that last one, I just wanted to clarify. So the one thing that the claims data were being used, in terms of adjusting the relative values, I heard you say that it was to reconcile the inpatient schedule of relative values with the outpatient schedule of relative values and was that based on claims or is that based on paid amount?

Just wanted to make sure I understood that. It's claims right? I mean the billed amount, not the paid amount.

Susan Knudson: Well, it's using the billed among the different services as (Gary) had explained to get the relativity among all the line-item detail. But then we used the pay amount to that paid-to-billed ratio to adjust amongst the different sectors of care.

Male: I see. So the, yes, that was my question. So between sectors, that's based on paid amount?

Susan Knudson: Yes.

Male: Okay, great. Thanks.

Male: So it didn't have anything to do with our views? So in other words, they know that they're spending if they're spending 40% of their cost on outpatient care...

Susan Knudson: I'm sorry?

Male: If they're spending 40% of their cost on outpatient care, then the outpatient care's relative values are going to have a, sort of, a 40 - account for 40% of the resource.

Susan Knudson: That's correct.

Male: Of resources in the end. I mean, that's what they're doing.

Male: Correct.

Bruce Steinwald: That sounds right although they can - is that right, Partners?

(Sally): Yes, that's correct.

Bruce Steinwald: Okay. Would anyone like to speak to the is(Sue) of reliability?

(Sally): I always like to make the statement, being within the insurance industry, that we are all assuming that administrative claims data is accurate.

Female: We're making the same assumption about clinical data as well.

(Sally): I guess, you know, when we're doing these analysis, it's just, you know, we might somehow has a caveat about how accurate these things are from a coding perspective. So it's a concern that people can, if they get too precise on their reliance on what this is telling us, you can really become misled, I guess I would say.

Bruce Steinwald: But that's going to be true for everything we look at, right?

(Sally): Absolutely, absolutely.

Female: As far as reliability metrics, it looked acceptable to me. I mean, there's a lot of different ways to measure reliability. But they were looking at sort of a boot-strapped hold against actual ((inaudible)) concurrent so...

(Sally): Yes.

Bruce Steinwald: Yes, I'm no expert but it looked reasonable.

(Sally): And I guess the only other is(Sue) in terms of something in the nature of claims payment (collect data) is that, people may or may not be aware of, the timeframe in terms of claims lags before you can do an evaluation on an entire, you know, set of data that apply to particular cases.

The hospital-based claims lags much longer than the professional, you know, pharmacies instantaneous. You know, so there's a lot of different timeframe things that need to be taken into account if you're looking at total care.

Bruce Steinwald: But they named only one time lag, three months, for all kinds of services, I think.

(Sally): Yes. And maybe they're doing okay with that but at least in my markets, and particularly if it's a complex case and it's, you know, a complex hospitalized case, that would not necessarily capture everything.

Bruce Steinwald: It wouldn't in Medicare for sure.

(Sally): Yes.

Female: This is commercial though. I think they have to bill by a certain point of time.

Susan Knudson: Yes, this is (Sue). You know most of our claims honestly are coming in electronically so we see them pretty quickly but we allow for the 90 day standard because that's what contractually we allow for. But we do see them coming in. And as it relates to the comment about the higher cost inpatient states, which might be associated with outliers, some of those may be hitting the truncation limit too. So a few things come into play.

Female: Yes.

Bruce Steinwald: What about 2b1 as also an element to reliability?

(Sally): I guess - this has to do with capturing the population and since this is a non-condition specific measure, I do think that some thought needs to be put in to looking at people that may be in a

population or in a panel that are inappropriately low utilizers because that represents an inappropriate use of resources and a risk.

Bruce Steinwald: Well we know they've excluded people who had no primary care doctor visits.

So that's the population in particular you're thinking of?

(Sally): Yes, and I can see within their system that this may be giving them all the information they need to feed back to providers on how they're utilizing services. But if this measure is applied to a larger population, these other is(Sue)s may become more significant. And also as they apply to disparities, sometimes the disparity marker is under utilization and I think that's coming up, so...

Bruce Steinwald: Good point. Anything else, any other comments on the two reliability sub-criteria? Then can we go on to validity testing?

The data, the sample, the methods, the results of testing?

Ashlie Wilbon: So this hearing committee could ask amongst yourselves whether the data and sample use is adequate for validity testing, discuss the analytical approach and then the findings themselves and whether you think that the testing demonstrates that it is measured the areas that, the measure set out, as described to measure and it distinguishing the different high and low costs of resource use.

Bruce Steinwald: I was satisfied but I'd like to hear from people who have more technical expertise than I do.

Female: Well it seems like they have a good sample size so...They had the Twins Cities Metro area, which is fairly large. And three years of measures and three years of claims...

Male: Sorry, you're very faint.

Female: Well I'm sorry. I said I think the sample size looks good. It's the Twins Cities Metro area, which is quite large and 19 different providers and 200 and some clinic sites so and the membership over the three year period was anywhere from 268,000 to 300,000, 3 - 303. So it seems to me that's certainly an adequate sample size.

I guess the one question I would have is what's the total population for Partners?

Susan Knudson: Well locally, we have about 700,000 members that's within this marketplace area locally. But that is Medicaid and Medicare as well and what we were looking at here was our commercial population who are attributed. Our non-user rate is around 9%.

Female: So it looks like you had about 50% of the population for commercial in your validity sample so that looks great.

Male: Does anyone want to comment on the correlations in the attached file that had correlations review and that data, if it all looked acceptable?

Particular on page 4 of that, since we're talking about the RUI, it said the ACG adjusted RUI was - 0.09 when it's at mean overall.

Bruce Steinwald: Well (what it works) is we're not hearing regular remarks. I'm, as I said, generally satisfied but because they're dealing with larger numbers than I think we're used to seeing except in some of the Medicare studies.

The correlations, I don't know, I have to evaluate them personally. What does other people think?

I mean it seems like there's, you know, overtime, a lot of stability here.

Female: I don't have any expertise to dive into it any deeper. It looks, you know, as you said, stable.

Bruce Steinwald: Alright can we move on to 2b.3?

(Sally): This is Sally. I just want to let you all know that we will be providing, in the near future, an in-depth summary of the testing findings. So when you retouch on the other measures, you can have an opportunity to have those - it'll be like a standardized assessment sheet to help make this conversation more robust as needed. And it'll be just focusing on testing approach and results.

Bruce Steinwald: Okay and we're good.

Exclusions and I think the excluding the subpopulation that has no primary care visits is an is(Sue) that was already raised and might not be much of an is(Sue) in the population actually studied, but for a broader scale, probably would be.

Ethan Halm: This is Ethan Halm. I agree that that's a generalizability concern. And in a way, you're excluding the sort of never-users and they also excluded the super high users by truncating them out.

Bruce Steinwald: Right and I guess they -- Partners, maybe you could respond to this -- you said you have a very group-oriented market but I wonder if there are physicians who are excluded from analysis because they aren't members of groups within your market. And how do you address

that problem or do you even see that as a problem?

Susan Knudson: We have not seen that as a problem and I might add, at least in our - we know from doing another benchmarking study just to understand how our own performance compares to national, regional and local benchmarks with a consulting firm that we have a relatively lower non-user rate. But I understand nationally, that's probably more like 17% from that study, if that fact helps. I realize it's just a one piece of information.

So and we see the bulk of our members do get attributed in this model to primary care. A smaller percent of those who are using see only a specialist, for example.

Male: And then you excluded those over 64.

Susan Knudson: Yes, this is a commercial model. Right now, we have not applied this to Medicare because we do not have the full suite of claims on our Medicare-covered population and so, we would be unable to do a full cost or resource use measure for those individuals.

Male: So would - is that impinge on this being applicable to a Medicare population?

Bruce Steinwald: Mm-hm.

Susan Knudson: Oh not necessarily because several will have full access to all that information.

Male: I'm sorry. I don't understand what you just said.

Male: I don't either.

Susan Knudson: So for example, as a plan providing coverage, we do not have hospital claims associated with the bulk of our membership because we have a cost product end, not a Medicare advantage where we'd be at full risk for everything. So therefore, we don't have those hospital claims. So I'd be unable to produce a full picture of resource use or total cost for the Medicare population.

Male: No, I understand what you said, that you didn't include the Medicare people because you don't have the data.

Susan Knudson: Okay.

Male: So my question is but you built a model here that's not been tested on Medicare enrollees so is that a limitation of the product -- that it's never been tested on Medicare enrollees?

(Sally): This is (Sally). Let me just - I'll give (Sue) a second to answer the question about limitation but as a reminder, for NQF endorsement, we would only endorse this measure for use in the commercial population. So if it's only been tested in the commercial population or if it'd only been tested in the Medicaid population, then that's where it would be endorsed for.

So that's something that is displayed to the users, what population this measure has been tested and is ready for standardized use.

Susan Knudson: Thank you (Sally). This is (Sue). Yes, exactly. That's what we've tested it on.

Bruce Steinwald: How small of a practice do you think that you can apply these measures to? The 600 patient criterion seems awfully large to me. Have you tried to apply these measures to small groups or even partnerships with individual practices?

Susan Knudson: You know, for what we've tested, it's been at the group level with 600 so we have not tested at the individual level with lower thresholds.

Male: And the reason - how we came up with the 600 was looking at the available quality measures and patient experience measures and how those samples lined up, just kind of what came with the 600.

Bruce Steinwald: Okay. Thank you.

Male: It's 600 group or 600 ((inaudible))?

Susan Knudson: Per group.

(Sally): This is (Sally). We have about ten minutes left. We want to give at least five minutes for the public to comment or provide input or ask questions. So I think we have time for one more question or thought and then we'll move it into public comment and then bring it back and adjourn for today with some heads-up on next steps to be expecting from NQF.

So Bruce, I think we have time for one more question and then I'll ask the operator to open it up.

Bruce Steinwald: Okay. Well, then it appears that we're not going to get through scientific acceptability. And then so we won't be able to address feasibility and usability.

But we will schedule another call for that?

(Sally): That's correct. We will be.

Bruce Steinwald: Okay. Well since we only have time for one more is (Sue), I'd like any steering committee member to put on the table anything that they think is really important that we haven't discussed on scientific acceptability.

Ashlie Wilbon: Yes. The only other thing and this may come up later, just had to do it at some point in time, we're to look at peer group averaging as a benchmark and how, you know, is that a good measure in all markets. This might be attributed - be utilized end.

I mean within a commercial payment scheme and network, it may work out okay. I'm just wondering how broadly applicable these measures might be utilized and whether that's - we'd have to do some serious thinking about how that will be used.

Bruce Steinwald: I agree. I mean, I think the fact that it's a very group, large group-oriented measure and most of the U.S. population is not served by a large group is a factor that we need to consider.

Ashlie Wilbon: Right. Or that there's - I think, compared to the Twins Cities, a lot wider variation not just in practice size but in probably some degree of quality follow-up. You know, it might be a little different using an average in a market with more scattered (servicing).

Bruce Steinwald: Well put. Any other steering committee member like to raise any is(Sue) before we turn it back to (Sally) and public comment?

William Rich: Yes. This is (Bill) Rich. You know, I spent many hours going through the algorithms of different groupers in the past and to be frank, there's some technical aspects that are way above my head and I'm being very honest. I had difficulty contributing to different parts and I think this is a learning process. It's our first one, but I'm being very frank and that some of the technical aspects were way beyond my ken.

Bruce Steinwald: You're not alone. I appreciate that, thank you.

William Rich: (They're just paying none of us).

David Redfearn: This is David. The one thing I would just comment is that it's - the methodology is enormously complex.

Bruce Steinwald: I agree David.

David Redfearn: And that's a concern I had in terms of not so much usage for professionals, but public use. I think that complexity kind of gets in the way.

Bruce Steinwald: Okay. (Sally), let's turn it back to you.

(Sally): Okay. Thank you. We have a couple things that Ashlie wants to touch bases with all of you on and then we're going to go straight to public comment.

Actually, well, let's go ahead and do public comment now and then I'll wrap it up with our final steps.

Ashlie Wilbon: Tom, could you please open up the call for public comment, questions or input at this time?

Operator: Yes ma'am. All lines are now open.

Ashlie Wilbon: Thank you. Any questions, input or comments from the listeners on the phone?

Okay. Well Tom, we can go ahead and I don't know how you - close the public comment portion of the conference call and then I'll talk to the steering committee members.

So I think we probably got about 5 - or maybe 3 minutes left, so just wanted to wrap up the call and give you guys some follow-up information and next steps.

So like I said at the beginning and we realized that we probably would not get through this entire measure, but I think it was probably to our advantage to just focus on one measure for this call. So and that's where - we're going to send out another survey for you guys to fill out with your availability for the next couple months. We'd like to get in at least one call before you come here for the in-person meeting.

If we're able to get maybe one or two calls in there to discuss the measures, hopefully we'll get through them before you come here to D.C. If not, they'll go in the agenda for the in-person meeting along with some of the cardiovascular diabetes measure. So, like this is a new process for us as it is for you so we're kind of taking things as they go, seeing how things go and trying to accommodate everyone with meetings and calls to get the work done in the timeframe we have allotted.

So with that said, I think even for the quality measures evaluations that I've heard, the committee made great strides today. So, it was very impressive work in an hour and a half so I'm very optimistic that we'll continue to get these measures evaluated in the discussions and questions out there to the surface.

(Sally): Absolutely. So on that note also, I would encourage everyone -- I realize everyone didn't get a chance to complete the survey - the actual survey, excuse me, SurveyMonkey Survey with the evaluation criteria in there. That link is still open. We were able to tally some of those results

before the call but we only had enough time to share them with the Co-Chairs.

Ashlie Wilbon: So what we'll do is if you would still, you know, try to submit your evaluation and we'll distribute those at various points in time and share with the group at an escorted format so that you can see what people have - type of comments people have submitted. Particularly after this discussion, there may be some things that came up that jogged or, you know, brought up other thoughts about different parts of the measure so we'll do that along with the availability survey and -

Bruce Steinwald: Is that sort of the final ranking that people are doing, is through that SurveyMonkey or is it something -

Ashlie Wilbon: I'm not - well, right now we're just at the point where we're trying to bring out any discussion points for the measure. We will, at the very end of the process, when you guys feel like you've been through all the criterion/criteria to a point where you're ready to vote on the overall measure. That will be when we'll ask you to give us your final ratings.

Bruce Steinwald: I see. So these are just kind of tentative or preliminary.

Ashlie Wilbon: Right. So we realize that as the discussion goes on, that people may be, you know, swayed in either, you know, one direction or the other and as you learn more about the measure, your ratings may change over time. But it just helps kind of share information with everyone and get an idea of where people's thought process is.

So that said, we'll keep HealthPartners and Ingenix in the loop as well so they will be informed of any calls that we schedule so that they can be on to answer questions. I think it's always very helpful to have the developer on the call to answer those technical questions.

So thank you (Sue) and your team for joining us today as well.

Susan Knudson: Thank you Ashlie. We appreciate the opportunity.

Ashlie Wilbon: If no one has any more questions, we'll go ahead and close the call for today. And if you come up with anything, questions or comments on the calls or the measures or anything, feel free to email (Sally) and I and we'll get back with you.

Paul Barnett: What's the status of - I'm sorry to have missed the last one, I guess on the twenty - a week ago anyway and it was going to be recorded. And that link doesn't seem to work still.

Ashlie Wilbon: Actually we had posted the - I'm sorry Paul, I was supposed to get back to you - we had posted the audio of that session. So for reason, we've been have some is(Sue)s with our webinar service so you were able to get the audio along with the slides, I think if you could listen to the audio and open the slide presentations, you should be able to follow along so I'll -

Paul Barnett: If I go to that site, nothing - I don't get any audio.

Ashlie Wilbon: It's a separate link, actually, that's posted on the Web site. It's an MP3. It's not actually - the link I sent you was from the webinar service, but we also have an MP3 that's just strictly an audio file.

Paul Barnett: I see and that's on the NQF site somewhere?

Ashlie Wilbon: Yes but I can send you the link to that. I'm making a note right now.

Paul Barnett: Thank you.

Bruce Steinwald: Alright, good bye.

Male: Thanks so much

Female: Thank you very much.

Ashlie Wilbon: Okay, thanks everyone.

Paul Barnett: Okay, bye bye.

Ashlie Wilbon: Bye.

END