NATIONAL QUALITY FORUM

Moderator: Sheila Crawford
May 9, 2014
11:00 a.m. ET

Operator: Welcome everyone. Today's webcast is about to begin. Please note today's call is being recorded, please stand by.

Karen Pace: Good morning everyone. This is Karen Pace here at NQF and thank you for joining the call. We're going to start with the roll call of the expert panel just to make sure we know who's here. And I'm going to just note that we have a separate line for the expert panel. So if for some reason, when we call your name and you're on but you're not on the speaking line, please press star zero to get the Operator's attention so that we get you into the right line.

So, Erin, will you do the roll call (inaudible) here?

Erin O'Rourke: Thanks Karen and I apologize in advance to anyone who's name I'll butcher.

Kevin Fiscella?

Kevin Fiscella: Yes.

Erin O'Rourke: David Nerenz?

David Nerenz: Here.

Erin O'Rourke: Jean Accius?

Jean Accius?

Alyce Adams?

Alyce Adams: Here.

Erin O'Rourke: Mary Barger?

Mary Barger: Here.

Erin O'Rourke: Susannah Bernheim?

Susannah Bernheim: Here.

Erin O'Rourke: Monica Bharel?

Monica Bharel: Hello, I'm here.

Hi, can you hear me?

Erin O'Rourke: Yes, thank you.

Mary Beth Callahan?

Mary Beth Callahan?

Lawrence Casalino?

Lawrence Casalino: Here.

Erin O'Rourke: Alyna Chien?

Alyna Chien?

Marshall Chin?

Marshall Chin: Here.

Erin O'Rourke: Mark Cohen?

Mark Cohen: Here.

Erin O'Rourke: Norbert Goldfield?

Norbert Goldfield: Here.

Erin O'Rourke: Nancy Garrett?

Nancy Garrett?

Nancy Garrett: Yes, I'm here. Can you hear me?

Erin O'Rourke: Yes.

Atul Grover?

David Hopkins?

David Hopkins: I'm here.

Erin O'Rourke: Dionne Jimenez?

Dionne Jimenez: I'm here.

Erin O'Rourke: Steven Lipstein?

Steven Lipstein: Here.

Erin O'Rourke: Eugene Nuccio?

Eugene Nuccio: Here.

Erin O'Rourke: Sean O'Brien?

Sean O'Brien: Here.

Erin O'Rourke: Pam Owens?

Pam Owens, are you on the line?

Erin O'Rourke: Ninez Ponce?

Ninez Ponce: Here.

Erin O'Rourke: Thu Quach.

Thu Quach: Here.

Erin O'Rourke: Tia Sawhney?

Tia Sawhney: Here.

Erin O'Rourke: Nancy Sugg?

Nancy Sugg: Here.

Erin O'Rourke: Rachel Werner?

Rachel Werner: Here.

Erin O'Rourke: Thank you everyone.

Mary Beth Callahan: Hi. This is Mary Beth Callahan, can you hear me?

Erin O'Rourke: Yes, thank you.

Mary Beth Callahan: OK good.

Karen Pace: So, I'll just do one last check. Jean Accius?

Alyna Chien?

And I think Atul Grover were not, and Pam Owens.

Female: Atul sent a note this morning.

Karen Pace: Yes.

OK. Thank you all for joining us. I'm going to do a quick process and agenda and then I will turn it over to Dave and Kevin for some opening remarks.

So you know, we' scheduled this time for four hours, just in case, then hopefully we'll end before that.

We will take a break at 1:00 Eastern time, 10 minutes just to give people a chance to get a little break in between, and we'll basically be working through the comments seen that we identified in the briefing memo and tried to work through some of those comments and discussed potential responses to them.

So, I will get into that more as we go along but for now, I'm going to first turn over to Kevin and then Dave to make some opening remarks.

Kevin Fiscella: I want to thank everybody for blocking out their schedules for this three-hour meeting. I know it's not easy. We all have very busy lives and I think as everybody has read the report, has gleaned lots of responses from many different parties and including some national press and I think our task today is to really begin to synthesize those comments and with our collective wisdom, come up with a balanced approach to those concerns.

I don't think either Dave nor I have a pre-conceived idea as to how the results of this discuss will turn out, but as before, we want to give everybody a chance to voice their own perspective and to – and if you haven't already bring any new data that hasn't been decided to the task.

Nancy Garrett: Kevin, this is Nancy. I heard you say the three-hour meeting. I have four hours blocked off.

Kevin Fiscella: Yes, four hours, I'm sorry, yes you're correct, four hours. Yes, it's exactly four hours. We're hoping we can – we can end before 3 but we have four hours blocked off. Thank you for the clarification.

Nancy Garrett: All right. We're ready to go. Thank you.

Karen Pace: OK, and Dave?

David Nerenz: I just want to be active around Kevin. Thanks to everyone, maybe it's a slightly different flavor, in looking to actively work, we've done so far, I think we should all take pride in the discussions we've had, the ideas we've generated, the nature of the draft report that's been created, and I think what that reflects is not only a lot of really smart and good input but also a very

high level of mutual respect, collegiality and openness listening to this, and these are issues that can get heated sometimes. I think our discussions have been really excellent, the in-person, the prior conference calls and I certainly – nothing but hope that we'll carry out forward as we go along.

I know as Kevin said, I think our charge at this point is to, you know, make whatever enhancements or modifications for the reports that we think we should base on the extent of comments that came in. Helen will tell us a little more about exactly, how ht process needs to work. But I – in looking at what we've done so far and what I think we can do by the time we get to the finish line, is very, very appreciative of all the time, the thought, the hard work that's gone in to where we are now. We look forward to even more as we move forward.

Karen Pace: OK, thank you. I asked Helen Burstin to, just – we had put some things in the briefing memo about NQF process and adjusting comments that I asked Helen to share some things with us about NQF process, so Helen?

Helen Burstin: OK, thanks everybody. So I do want to also just extend my thanks to the entire committee but also in particular to Dave and Kevin who by far in my seven years at NQF have been the hardest working chairs and most engaged I've ever worked with. So really thank you.

And also a special thanks to Karen obviously for her amazing stewardship of this work. I will tell you several, very high prominent policy people have sent some comments to me that are along the lines of you know, this is probably the most important issue in CAUTI measurement, one of the most important debate in today's pay-per-performance.

So there's no question, this is a timely and really important piece of work and just a few words about consensus, and I really do want to emphasize that as a consensus-based organization, our goal here is actually to find common ground, see if we can actually review these comments, find a place where we can actually find some resolution, because I think this is not the kind of issue where I think we want to wind up having a deeply polarized membership.

I think this is hopefully a place we can actually find that common ground through the four hours of our time together today.

A little bit about consensus and what it means and Karen included some of that in the memo for you but the definition of consensus that we used as a standard setting organization, as dictated by the – on the circular is that it is general agreement, but not unanimity. And I think that's really important. But general agreement is also important. We don't want to wind up here having significant disagreements as a result of this.

It goes on to further state that our goal here, to really be attempting to resolve objections by interested parties, and we need to fully consider each objection – each comment and then tell everybody why and how those objections or concerns were addressed. With 670 comments, that is not an insignificant task but obviously, many of them fit into similar categories and I think (inaudible) does a great job of laying out those ideas.

But just in terms of Next Steps, while there are so many comments, 670 also a record for NQF, you know, overwhelmingly, in terms of the numbers, they are overwhelmingly positive that they'd also overwhelmingly heavily from, you know, two stakeholders, providers and particular, in health professionals. But we need to consider the comments the across the live variety of stakeholders and there were some consumers and purchasers who are positive, actually it's three in particular, but there were also some that were negative.

But as this report moves on to the next steps in our consensus process, to the feedback, and in to the NQF board, those are groups that are consumer and purchaser majority groups. So again, I can't emphasize enough how important it is for us today to try to consider the comments, consider whether there opportunities to in fact, bring people together and see if we can find common ground because we'd very much like to have this report finalizing out there in early summer. I think a lot of people are looking to this and hoping that it will help really address what I think, you know, act really prescribed by many as being something that tends to be such an important block in our ability to do better in terms of performance measure and to start these reductions.

So with that, I will turn it back over to Dave and Kevin.

Karen Pace: This is Karen. I'm going to jump in here for just one second and just say that I think before we get in to the specific comments and themes that we wanted to give the – each of the panel members the opportunity to make some brief comments about your general impressions or any specific issue that you think is, you know, really important for us to discuss today. And I'll start with Kevin and Dave there.

So Kevin and Dave, if you want to start about and then I can go through the list so that – I know it's hard. We're at a disadvantage because we can't see each other.

Kevin, do you have anything you want to say in general or general impressions about the comments, or specific issues, or …

Kevin Fiscella: Yes. I don't really have too much to say. I think that the issues have been laid out and some are issues of clarification and response in terms of evidence. Others relate to issues of essentially unforeseen or potentially unintended consequences of either not adjusting or adjusting and thinking about potential mechanisms to monitor for example, those unintended consequences. And some of at came up both in the public comments as well as comments in some of the webinar. So I think we need to think about both pieces.

Karen Pace: OK, thank you. Dave Nerenz?

David Nerenz: I think just in broad framing of I think the agenda that we have in front of us in the memo that it's going to guide us through what I can see the main issues. You know, when you look at the spectrum of comments, it's reassuring and you feel good when you possess number then say, you got it right and thank you for the work you've done, but we really have to focus on those that express concern or objection among, you know, it seems like there are three themes that come up in essentially all of them and there are themes actually we all had among ourselves when we started this process.

If there is adjustment, does that heavily affect masking disparities, does it have the effect of creating a low standard of care for providers to plan their

sort of disadvantage patients or client members and the lower incentives? I think if I was asked for a, you know, what's the main critical comment? It would be those three issues, but I think we have those PDF in front of us already in terms of the agenda so I'm sure we'll speak to those.

Karen Pace: OK, thank you. At Alyce Adams, do you have any brief comments that you'd like to make to start us out?

Alyce Adams: No. I think I completely agree about the comments that were just made about the fact that these are all issues that came up during our conversation so certainly none of it is a complete surprise. I think it would be nice if we could come to a consensus where we were able to minimize both the risks associated with adjustment and the risks associated with not adjusting. I don't know if such a nice common ground exists, but I'd like to get there. And I guess that's it.

Karen Pace: OK, thank you. Mary Barger?

Mary Barger: I don't have any additional things to add. I would just echo what most had just said.

Karen Pace: OK, thank you. Susannah Bernheim?

Susannah Bernheim: Hi, yes. I think that from my perspective, the comments particularly coming from the consumer and purchaser group point to a need for us to help – I think if we can help to operationalize a little bit more the recommendations we have which is not a simple task, it will help people sort of see how this might play out and that that's going to help people because I think the recommendations are meant to carry a sense that there are times when you would want to risk adjust and times when you wouldn't.

And I think the concerns from that group are that the report found more sort of totally pro-risk adjustment than I think the consensus and the committee is. So I think if we can help just clarify better kind of how these decisions will be made even if we can't explain it perfectly, people will feel like there's a task for measures to be individualized and I think that will get us a long way towards responding to these comments.

Karen Pace: OK, thank you. Monica Bharel?

Monica Bharel: I agree with everything that's been said particularly this issue around – we have such a nuanced discussion and are there ways that we enhance the report itself? I think a lot of us have consensus there on the issues but the way report we – maybe there need to be some clarification and I can make suggestion as we go around those nuances.

And particularly for the comment in general, the comments that came through the concept of risk adjustment that we're already doing for many clinical issues, we've talked about this several times such as something like diabetes and how these current risk adjustments that we're talking about parallel the clinical ones. I think we can do some clarification there as well.

Karen Pace: OK, thank you. Mary Beth Callahan?

Mary Beth Callahan: Can you hear me?

Karen Pace: Yes.

Mary Beth Callahan: OK. I think that this is excellent work. And I think that in the memos that went out ahead of time, some of the responses as potential clarification are very good and I know we'll discuss those later. My greatest concern is of course the things I've seen in Texas happening with Medicaid managed care and the problems we've had for patients getting access to doctors because of what happened with Medicaid managed care. So I'm kind of on defense with all of this but I think the report is very good.

Karen Pace: All right, thank you. And Larry Casalino?

Lawrence Casalino: Yes. I think first of all, I thank you for the memo. I think it was extremely useful and helped me in reviewing the comments. I thought the comments were really good. I think that – I am talking about the work, I'm taking about the critical ones now specifically. You know, they raised a lot of issues but I hope that we'll be very careful not to kind of get caught up in relatively micro issues and being tired or not have enough time to deal with I

think the biggest single criticism operationally which is the idea of presenting one single number adjusted for SES.

I think that's really the crux of what's going on. I hope we'll spend a good time on that. I actually think there's a – that there's fairly a simple way of responding to that that it could satisfy everybody. And that's it.

Karen Pace: Thank you. And Marshall Chin?

Marshall Chin: I think we look at recommendation one, how we divide it into purpose of accountability, risk adjusting by SES and in the second bullet about – for identifying – do some disparities, have stratified measures. They're never going to attempt to basically have – give the (fairness) issue with risk adjustment, but also have an act that kind of – that aims up at (sunny line) disparities and encouraging solutions.

Looking at the comments, people tend to focus on the first part of recommendation but not on the part we've mentioned about identifying disparities and stratifying data. And part of that is to have – potentially, we need – we need to have more in the report about falling that line of thought so that people concerned about welts or we met disparities, are we giving a pass for these various issues?

Part of that response is clarifying how that second bullet would not have or reducing disparities. So best overall and then the current recommendation one tries to do both but the second part about actually, the disparity solutions, perhaps we can work on how we make that argument clear with that discussion clearer and that may help with some of the – I guess (over gusto) people – critics have – they're taken away from this of not addressing disparities.

Karen Pace: OK, thank you. Mark Cohen?

Mark Cohen: I have nothing to add right now.

Karen Pace: OK. Norbert Godlfield?

Norbert Godlfield: Can you hear me OK?

Karen Pace: Yes.

Norbert Godlfield: So I will make two comments. Number one, to put it in a phone-appropriate manner, I was quite interested to see that the report hit the New York Times. And the headline, I believe addressed head-on the issue that has always been the 800-pound gorilla even though there has been some, you know, statements from time to time that in fact, payment is not being addressed. In fact, you know, it's all about payments at the end of the day.

And so I just want to highlight as my first comment that they are looking at historical perspective, (PPS) would never, never – have been implemented in 1982, if there had not been disproportion of chair policy. So that's my first comment.

My second comment is that I believe has already been stated and I think the – in fact, we had researchers neither agree with the MedPAC approach even though we do a lot of work with MedPAC, nor do we believe in (necessarily) one score approach. So we would suggest and I would suggest letting 1,000 flowers bloom because there's a lot that we don't know that we really – that SES must be addressed, but there are many different ways to doing it and there's still a lot we don't know, and I certainly (inaudible), it's very efficient (inaudible).

Karen Pace: OK. Nancy Garrett?

Nancy Garrett: Yes. So I think to me, one of the most interesting things about today is how do you take all of these comments? There are so many that we got. I mean, it sounds like it was kind of a world record for NQF in terms of the number of comments. And how do you balance what Helen was talking about the idea of consensus around the stakeholders but make sure we're moving forward with, you know, the recommendation and our committee really feels we can stand behind? So that balancing and how to take into account the comments and how to – but still have a balanced perspective. It's going to be – I think it's going to be tricky.

So we'll be looking to the NQF staff for guidance I think on how to do that. I'm a little bit worried about – that we might take some of the comments not in support and give them more weight than they really should have given that we had so many in support. So that will be an interesting thing I think as we go forward.

Karen Pace: OK, thank you. David Hopkins?

David Hopkins: Hi. Good morning everyone. Well as you know, I'm the one that represents the consumers and purchasers here, not that we all don't in one way or another but that's my specific role. And so as been noted, we're the skeptics, not yet convinced I think is the best way to portray it. Dave Nerenz did a wonderful job a few minutes ago of summarizing very briefly the major concerns about unintended consequences.

We were looking for more evidence of the kind that NQF usually employs and make decisions for doing things one way or the other still looking for that. I think that it's one of the things we could do today and it's already been mentioned I think by Susannah is to reach more specificity on the recommendation. So if it's appropriate to include SES in risk adjustment, can we get a little bit more specific about under what circumstances or what kinds of measures, what SES factors, what data and so on and so forth to make it a lot more practical?

Thirdly, I would hope that we could try to separate payment issues from measurement issues, not an easy thing to do in this day and age. I think our focus here is on getting the measurement right. And there are plenty of places where we can try to work on payment and I think they're – we're all with you. I'm thinking the safety net providers here. We recognize your burden and the resources not being sufficient and would be just as happy to get involved with you in some way to do what we can to help improve on that.

And then finally, to pick up on what I think heard Larry say maybe we can come up with some kind of compromised solution this morning that will satisfy – will serve as the next step as a way to move forward. I'm hoping for that and I have some thoughts.

Karen Pace: OK. And just a comment – our next person is Dionne Jimenez and we actually had two other representatives on the panel from consumer and purchaser organizations. Dionne is from Service Employees International and Jean Accius with AARP. But I think – Dionne, you're on.

Dionne Jimenez: Oh sure. Well thank you for that Karen because I was going to mention – yes exactly, (we also) represent sort of the consumer and purchaser voice. And, you know, we definitely were supportive of the report and the recommendations and we do have a few suggestions of how to approve it. And I do agree with kind of what Susannah and David was talking about, if we can try to get a little more specific. And with what Nancy was saying in terms of if there's a way of how we can try to reach consensus because I think our goal as a group is we want this report to be adopted and hopefully, there's a way that we can find a way to appropriately address some of the concerns that were raised without necessarily getting away too far from our recommendation.

And, you know, I wish we had a longer period of time to have – not in person because I think over two days, you're only able to accomplish so much. And I think, you know, some of the calls for a lot of additional evidence and other things, I think it's hard to accomplish that in sort of the format that we've been looking with as a committee or as a technical expert panel. And hopefully, you know, I think we can – hopefully, we can adopt sort of a longer term committee that could address some more of these specific issues.

Karen Pace: OK, thank you. Steven Lipstein?

Steven Lipstein: Yes. Good morning and thank you again for all of the leadership that Kevin, and David, and Karen have provided. This is kind of a very important undertaking. One of the things that I did once I saw all the comments from all of the people who submitted comments was I knew a few of them personally. So I wanted to better understand the concerns of those who had voiced opposition in terms of trying to learn whether people were open to receiving additional evidence and would be willing to change their minds or change their opposition if they had additional evidence because several had cited the fact that they thought the evidence was – and I used the word premature, the

evidence was premature to support the overall recommendations for sociodemographic risk adjustment. And I was trying to understand how much additional gestation would be required to take our recommendations from a place being premature to being full-term.

And I think towards that end, I think what David and Susannah said earlier is very important in terms of the opponents to our recommendations being skeptical and not yet convinced. And so I kind of was able to identify four themes that were common among that skepticism. One was that, you know, a concern – and I'll use the readmissions rate example, that providers who serve low-income communities may have higher readmission rates because they don't do a very good job of providing discharge summaries, or discharge planning, or patient education, or transition of care management.

So how do we make sure that sociodemographic risk adjustment doesn't embed substandard care? And so, you know, so I think that's one thing that we need to address and I actually thought the article that I sent around yesterday that was the Henry Ford Hospital example showed that even with standard and consistent discharge planning, discharge summaries, transition of care management, you still had highly differential readmission rates when people were discharged to high-poverty communities. So I think that's one theme that we heard.

The second theme was that providers who serve low-income or disadvantaged communities will continue to do more research and to continue to improve their performance even in the presence of sociodemographic risk adjustment. And so they don't want – so the kind of the skeptics would say, you know, "If we sociodemographic risk adjust and it takes away the motivation to improve, that's not a good thing." So I think that's a second theme we need to address.

The third one is this whole issue of disparities need to continue to be visible. And this is where we've had our discussions about how you do sociodemographic risk adjustment without making disparities invisible either to patients, or providers, or to payers, or the consumer community broadly defined.

And then the fourth theme which is an interesting one that I think we need to address is this concern that we were making it too difficult on the measure developers. And, you know, I raised the issue in – what I said around is in making it easier on the developers, we make it more difficult on the providers of services to disadvantaged communities, I opt for making it easier on the providers and more difficult on the measure developers but I do think that that's a concern that was expressed by the people who were in opposition.

So again, those were the four themes that I think I learned about – in both reviewing the comments that were in opposition to our recommendations as well as communicating with a few of the people who voiced those – that opposition who I know personally. So thank you.

Karen Pace: OK, thank you. Gene Nuccio?

Eugene Nuccio: Yes, good morning. I have just two comments. One is that I think we as a community who picked on a great job but I think we also need to be realistic about what risk adjustment can and cannot do with regard to creating an equivalent playing field for similar healthcare providers. So I think it's sort of realistic approach of where we are today, what the possibility of adding the socioeconomic status and demographic variables would or would not add is something that we need to keep in the back of our minds.

And just to provide a sort of kind of argument to my friend Steve, I don't think developers see it as a burden. I think developers don't want to be restricted or be told that it's – I have a prescription provided to us in terms of how we go about developing a risk adjustment model because there are many different measures that need to be risk adjusted, some of which are better risk adjusted and using an indirect methods, some using linear method. And I think that develops don't want to be put in a box but allow us to be creative in terms of how we address this issue of providing an equivalent playing field when we make comparisons among providers. So those are my two comments.

Karen Pace: OK. Sean O'Brien?

Sean O'Brien: No comments right now.

Karen Pace: OK. Ninez Ponce – Ponce? I'm sorry.

Ninez Ponce: Did you skipped over Pam Owens?

Karen Pace: I didn't know Pam was on. Is Pam there?

Pam Owens: Yes, I'm here. Sorry.

Karen Pace: Oh, OK Pam. Thank you.

Pam Owens: Coming from a business meeting. So actually, very much appreciate what the members have already said. The part I'd like to back to because I think it's a nuanced discussion is so how do we need to look at the language again in the sense from whether the comments are coming? But the nuances are – that I found to be really fruitful in-person meeting. Maybe that didn't come across as strongly but it is a nuance discussion because it seems like the comments that I read were – read it as more black and white than I think we – than I left the meeting feeling like our conversation was. I felt like we had a lot of caveats and sort of – it really wasn't black and white.

Karen Pace: All right, thank you.

David Nerenz: Can I just quickly respond?

Karen Pace: Yes.

David Nerenz: Dave here, I didn't want to interrupt. That's really important Pam, I appreciate that. I know several comments had been made like that. If there's a reason, it may have been that since we're – one of our main recommendations is different from the current status quo, we're probably felt the need to argue that particular change very strongly, and if so, perhaps want some nuance.

So as we go through our discussion for the rest of today, if we can highlight those areas of nuance that people felt were in the in-person meeting, lost in the report should be brought back. Let's make sure we do that.

Karen Pace: OK. Ninez?

Ninez Ponce: I think that – I thought that the comments were very favorable and even the ones that were not – gave significant recommendations. So I don't, I mean, I was a little bit anxious about this meeting but I think that it's not going to be such a hard task. One thing is accommodation, the easy thing is the list of adjusters that some stakeholders suggested, you know, adding from the suggested list.

And (inaudible), I feel that the group, at one point, it was a turn to try to separate the goal of, you know, performance and reducing disparities. And I think that two should still be aligned and I was surprised and appreciate (the health) insurance plan emphasize the need to talk about disparity.

Karen Pace: OK, thank you. Thu Quach?

Thu Quach: Yes. So I was really pleased with the number that the gap and the diversity of the commenters and their comments and I was like reading through many of them – I agree with what's been said. The one thing I want to underscore and highlight is this concern around the data burden and I know that in some of the questions that we had, we're looking to use existing data. And I would really want some more discussion around that.

I think in D.C., when we had sort of the breakout group and I was part of the group where we discussed about the data and the variables we want to use, I really always felt like we needed to extend that discussion and really evaluate the different types of variables that under consideration. Some of the commenters added a few others that I thought were really good but at some point, looking to what's available versus what really needs to be improved, I think that is going to be key for us in terms of the data quality and how precise and how accurate it is in terms of making or breaking the risk adjustment method.

Karen Pace: Thank you. OK, Tia?

Tia Sawhney: Yes. First of all, I was very, very impressed by the number of positive comments as a percentage. In talking to psychology, people often are more motivated to put in their complaints than their praise and yet, we got overwhelmingly positive comments and I think that's a very good thing. And

that's not to discount the ones that were negative, there's learning in those and we need to take them very seriously.

This isn't unknown, this is new. And as I said, it's incredible that we got so much support in my mind. I would hesitate to move to get too prescriptive because this is going to grow and evolve and change over future years. So the more we can do establish a good framework is more important to me than drilling down to prescriptive elements.

And along those lines, any measurement, any – exist within a system. I mean – and what is a good measurement for one purpose is not a good measurement for another purpose. And to respond to some of those comments that advised us not to give advice on how measurements get used and how they get implemented, I think that would be neglectful.

Giving advice is different than prescribing, but to the extent that we have knowledge about the inherent weaknesses or limitations of a measurement, I think we're duly bound to share them. So …

Karen Pace: OK, thank you. Nancy Sugg?

Nancy Sugg: Hi, good morning. So I just want to go back to the three things that David brought up at the beginning. And as I think through those and I look at masking disparities and I think we can have some very good answers to that, some ways to maybe reassure people. The – lowering the incentive to improve I think really comes down to what you use to incentivize improvement. So I think that there are some fairly easy frameworks to put that in.

The one I'm most concerned about is the perception that we're accepting lower standards for people of lower socioeconomic status and I'm worried about that because much of it is perception. But I think we do have to very thoughtfully think through how we want to give reassurance to that perception because I know sitting around the table with you all, none of us want that to happen. So I think that that's the place where I'm not as sure how we want to address that but I think it's very important.

Karen Pace: All right, thank you. Rachel Werner?

Rachel Werner: Yes. So I guess I'm just going to echo a lot of what some said already but as I read through the memo that's put together summarizing the comment and the comment themselves, I was primarily struck by the number of positive comments and I was – thought that it would be overwhelming – the comments were generally overwhelmingly positive.

And I was also struck with the fact that many other negative comments I think were around issues that we discussed at length in our in-person meeting and subsequently and that there's a lot of nuance required in talking about those issues and so I got a little bit of lost in that report. And so some clarification could help a lot in addressing some of those issues.

Karen Pace: All right, thank you. And did I miss anyone?

(Off-mike)

Karen Pace: OK …

Susannah Bernheim: Karen, can I ask one thing. Karen, can I just ask one thing, this is Susannah?

Karen Pace: Yes.

Susannah Bernheim: Following on what Rachel said and actually a number of other people's comments, it strikes me and I don't know if this is helpful but some of the polarization of the commenters and some of what I think they may have read in the report are as if this report was weighing in on should we or shouldn't we risk adjust threat, yes?

And I think if our conversation is about sort of should we or shouldn't we, it's not as productive as if we can try to focus on when should we and when shouldn't we. So getting guidance to the – and there may be people who feel like the answer when should we is always but I don't think that that's what the consensus of the committee is. And I think that framing might reassure the people who are concerned because somebody said this, I think it was read as

being more polarizing than it was intended to be. And so I'm wondering if that shift would help our conversation and our edits.

Karen Pace: OK, thank you. You know, and I think that is certainly a good comment and something for us to consider and I appreciate that. So what I'm going to do is just move on to this quick summary of the comments. Actually I think we'll – let's go to the memo. We're going to be looking at the memo and then we're going to start working through the themes and having some discussion. Dave and Kevin will help facilitate that discussion. And then we'll continue to see where we end up today.

So I think basically, everyone has read the summary of accounts to the comments and we've all commented that it was a large number and a record for NQF. And I think it just speaks to the importance of this topic and the engagement and certainly it starts with this expert panel and again, I appreciate the engagement with (inaudible) move down.

And I think we've already talked about some of the balance of the stakeholder comments and that's one of the things that we want to address in our consensus process. And certainly to keep in mind as we think through these comments and potential resolution or responses to those comments. So let's go on and move on down to the themes.

And we're Screen Sharing and we realize that there's a little bit of a lag between what we do here and what you see on your screen. We are basically going to be using the memo as our kind of agenda at least to start with. So the third theme that we wanted to focus on and many of you have already commented on is the comments related to masking disparities or masking quality problems and potentially accepting different standards of care for patients with different circumstances.

And I think one of the things that is noteworthy is that commenters agreed with the recommendation that stratification was necessary to identify disparities. However, some commenters did object to sociodemographic adjustment for purposes of public reporting and paper performance.

Expressed concerns, these things that we've already as mentioned, talked about during the meeting and subsequent calls.

And I think that – but there were other commenters that noted that analyses that would be needed to include adjustment for sociodemographic factors would highlight where there are disparities and actually provide more information to bring to the issue of identifying and working on disparities. Some commenters suggested that both the sociodemographic adjusted and stratified data be publicly reported.

So that's in a nutshell and we just identified some questions for you to be thinking about and then I'm going to turn this over to all of you for discussion. But again, should the recommendations regarding sociodemographic adjustment be modified in any way, are there ways to resolve or mitigate concerns about the sociodemographic adjustment and masking?

So for example, monitor for increasing disparities, incremental approaches, more transparency, should the disparities recommendation, the second part, recommendation one actually be separate so that it can be adequately identified and perhaps even strengthened? And Kevin had provided some additional suggestion in the appendix.

And then I think something that came up in some of your comments, are there additional ways to explain or demonstrate what a risk adjustment does and doesn't do? So from there, I'm going to turn it over to Kevin to help facilitate a discussion of some ideas of how we can address and respond or hopefully resolve some of the concerns that were expressed. Kevin?

Kevin Fiscella: Thanks, Karen. So I don't know that there's a perfect way to address this complex problem but I think one way would be to begin with the first bolded which is that we're (bolded) masking disparities, masking quality of problems under different standards. And starting with the masking disparities and thinking about our – the first part of recommendation, recommendation 1-A, and then I'll – what we think a response to the, you know, specifically to the masking issue should be.

I think it's going to be easier if we – although they're clearly interrelated, perhaps to take them one at a time since some of the public comments actually – I listed them separately.

Lawrence Casalino: David, this is Larry Casalino. Should I jump in or – how do you want to proceed?

Karen Pace: Yes …

Kevin Fiscella: Yes. Yes let's …

Lawrence Casalino: So OK.

Karen Pace: Just identify yourself when you make a comment. Thank you.

Lawrence Casalino: Yes. It's Larry Casalino from Weill Cornell Medical College. You know, we had some discussion – a fair amount actually during our in-person meetings about whether it made sense to report one risk adjusted measure. Let's just say, in cases where it seems appropriate that SES is important and we'll talk about later, maybe if we need to clarify when that might be.

But assuming that the – it looks like SES is important, you know, we debated whether to report one – whether the measure should be one risk adjusted measure, just SES, or an unadjusted measure, it might be clinically adjusted but not adjusted for SES, and an adjusted measure for SES. And we had I think a relatively brief discussion about that and that, you know, maybe it would be too complicated and too confusing to report both.

And, you know, looking aback, I was a little uneasy with that at that time. I have to say, in looking back at it and looking at the comments, I actually think that I think we would meet the most serious concerns that I think were expressed quite well if we said, you know, if SES should be addressed, then the measure should have two components. They'd be one, unadjusted for SES and one, adjusted for SES and they'd both be reported.

I'm not going to get into the comparing among straighter right now, I think that's a separate discussion. I think it's a good one. It's a good way to look at

disparities in a more detailed way for quality improvement. But just the adjusted and unadjusted, I really don't think that it's too complicated for public reporting, it's two numbers, they're easily explained and I don't think it's too complicated for payment incentives either.

I think one of our problems and I'm going to shut up a minute here not to go on a too much length. I think one problem is, you know, our charts with these two is considered measurement issues and not how the measures would be used. But in this area, it's very, very hard to do that as we see repeatedly.

I think if there is a concern about how two measures would be used, I think it's actually fairly obvious and would go a long way toward meeting the concerns of those who are concerned about what will happen to providers who take care of disadvantaged patients and also, the concerns of those people concerned about masking disparities and giving providers a pass and not improving care for disadvantaged patients.

And that is simply you have the two numbers, adjusted and unadjusted, and then you pay based on some blend of the two and probably based on some blend of the two and maybe an improvement as well. And maybe over time, perhaps a long time, you shift the blend more and more toward the unadjusted score.

This is all up for debate and I'm not sure that NQF needs to recommend that. But I just don't see that, it's that complicated to have the two numbers and I think it would solve a lot of problems.

David Hopkins: This is David Hopkins. Can – may I respond?

Lawrence Casalino: Yes, yes.

Karen Pace: Yes.

David Hopkins: Larry, that's exactly the compromise solutions that I have thought of. So I just want to say I totally support that. I think that it is the best answer to those of us who may remain somewhat skeptical and allows for, you know, any variety of solutions in the future on payment issues, public reporting issues and so

forth. And I'm always a little hesitant to use the word public reporting because we often think of, oh gosh, this is going to be confusing when we put it out on the web for our consumers. So I prefer to think of just reporting to whoever is appropriate.

So I just second Larry's proposal there and I hope maybe that we could agree on that. That would be a huge step forward and I think would quiet down the skeptics quite a bit.

Kevin Fiscella: And this is Kevin. I'm going to jump in here quickly. I wonder whether we should have some focused discussion specifically on this very important issue. We'll get back to the broader issue but, you know, having – and now that you've raised this Larry and David, you have responded, I think to have some broader discussion about this key point which I think does relate very much to one of the criticisms at this point I think would be helpful. And then we'll come back to the broader issue once we had more discussion about the issue of whether or not to include adjusted measures that are unadjusted for SES along with those that are fully adjusted.

Ninez Ponce: Kevin, this is Ninez. Hello, this is Ninez and I absolutely endorse this. In fact, I think I endorsed this from the very beginning. I think the way you unmasked disparities by having the adjusted and unadjusted and you see how it moved and I know our workgroup and what Susannah presented, you know, it was, you know, there would be – with those two adjusted and non-adjusted, you'd see the, you know, the room for improvement but also, that there are providers that are doing the best they can per se on a more underserved community.

So I, you know, I'm glad that it put out and I hope it stay as a response.

Male: This is …

Norbert Godlfield: This is Norbert Godlfield. Can you hear me OK?

David Hopkins: Yes.

Norbert Godlfield: So I just have two comments and, you know, this may be the path that we'll go down but I just want to highlight just from a developer's point of view and both Susannah and I in a positive way disagree but I think it highlights, you know, the issue which is to say, I said in my presentation, is that for me the poster child was homelessness and that should be incorporated to risk adjustment and a very close extension to chronically mentally ill. And although they are just – take an example, readmission

So I just want to highlight that there are going to be many measures out there if you separate out the risk adjustment from the socioeconomic. But clearly, I just want to hammer on the fact that there is not a bright line between the two and in fact, develops can differ.

And so what I would suggest as we go forward, if we go down the path that has been suggested by Larry and David, I think it's very important to really emphasize that the developer needs to be – to demonstrate a significant effort they have incorporated into the risk adjustment mechanism as much that reflect those economic status is possible that tie to risk adjustment, an example again being chronically mentally ill in (inaudible) where again, with no – with all, you know, probably (inaudible) developers can disagree and in fact, do disagree.

Kevin Fiscella: Norbert, do you have a petition on the question of whether to include both?

Norbert Godlfield: On the question of including both measures?

Kevin Fiscella: No, (both) adjusted and unadjusted.

Norbert Godlfield: Well again, I think the – my perspective at the end of the day, I believe that initially, so I'm going to temporize by saying that for initial efforts, it should be combined together. I've had no problems over time when the two be – to be separate because in fact, it's exactly the reason that we all support. But initially, as I believe, they should be combined together.

But if that can't go, I just believe also that there has to be a very diligent effort to incorporate into the risk adjustment mechanism as many aspects (inaudible).

Jean Accius: This is Jean. Just a point of clarification, when you're speaking of unadjusted and then adjusted, are you really speaking about raw values or are you speaking about values that had been adjusted for clinical patient conditions, and then the adjusted is the adjustment that occurs if a socioeconomic or a demographic variable is added into the prediction model?

Lawrence Casalino: Yes. This is Larry. Since I put it out, I guess I'll respond. So when I say unadjusted, I mean, unadjusted for SES, and I think that's what others are meaning too. The unadjusted would be adjusted for clinical factors in cases where it seemed appropriate to adjust for clinical factors. And this is already something that's been addressed I think well by NQF and NQF measure developers.

So to me, unadjusted is adjusted for clinical if appropriate but isn't adjusted for SES. Adjusted is adjusted for clinical if appropriate plus SES.

Tia Sawhney: This is Tia Sawhney and my apologies for getting technical here but this is an important technical point. In healthcare, everything is correlated with everything. So SES is in fact, correlated with clinical. It has its own distinct value but it is correlated. So if we're doing – say for example, risk adjustment is a multivariate linear regression, we could calculate the best – optimize the adjustment factors using just purely clinical.

But as soon as we put SES in, the clinical – the factors related to the clinical characteristics will change. So if we take SES in and out, do we re-opt – if we have SES in and then we say do it without, does without mean just removing the SES factor without re-optimizing the clinical factors, so does it mean re-optimizing the clinical factors?

Susannah Bernheim: Tia, this is Susannah. I think that you're right that there are real complexities to this but – and I suspect that we can't probably get into the details of them but I think you could do that either way. And obviously, one is more burdensome but on a higher level, I want to support the suggestion that we indicate and report that there will be times when the best solution is to provide those pieces of information and I think it is a great step towards transparency.

And as Larry said, it's not going to always be a very comfortable thing to have the two numbers out there but it is going to resolve a lot of the concerns people have. It's going to make it really clear what happens when you do risk adjustment, how big of a difference it makes, how much it's going to help you understand more about a provider.

And so in my framework of not – should we or shouldn't we but when should we, I think it would be a great thing to say and I don't have the right words exactly but just say there are going to be certain measures under which case, we always use the example of (CLABSI), there's no need to risk adjust (inaudible) so you would just present a measure that did not included (inaudible).

There are going to be other measures, you know, right now, it happens with HCAP where the risk adjustment for education factor is critical and should be presented just as a single number with the education level adjusted in. And there are going to be measures where it is more controversial and we recommend at least in the early stages that developers present both a risk adjusted for SES and without measure. And that's going to – I think it's going to help you but I wasn't as much of a fan in the meeting, but I really agree today.

Tia Sawhney: OK. So dig a little bit deeper there in what you just said in that, you – I think you just outlined three alternative paths in my mind because – or you added two to what we were previously discussing. In that – do we – or is there a recommendation that they should always be presented in parallel or they may be presented in a parallel that we would rather they be presented in parallel? Because for sure, we've always – they always – I would say that even our current recommendation doesn't preclude them being presented in parallel.

Lawrence Casalino: So this is Larry again, if I can jump in? Again, I think that was Susannah's comment, illustrates to how difficult it is to separate the measurement from the use, right? And so when we say presented, we're kind of presuming for accountability purposes. And I guess, I mean, I agree that when relevant – when SES has deemed important that both should be presented but, I mean, I

think we can – at a very basic level, we can just say – as we said I think that the measure developer should explain why they think SES is or is not important in this measure and why the measure should or should not be adjusted.

If there is no good reason to think it shouldn't be adjusted, then the measure would require two calculations, one of the adjusted, one of the unadjusted period. It's another step then to say how these would be used for accountability and we may want to, you know, provide some guidance on that. I think it's fairly obvious actually. But I think, you know, we're kind of getting into the other discussion of when but I would say if it's something where the patient has to do something, then SES is probably – it should probably – measure probably should be adjusted for SES. It would be a strong burden on the developer to show why it shouldn't be.

So again, Susannah gave the example of CLABSI. And I think we all agree the patient really doesn't have to do anything for that, it shouldn't be adjusted for SES. You start to look at readmissions and (demography) rates or many, many process and outcome measures where if they do depend to some extent on what the patient does, those I think unless the developer give a very good reason should be SES-suggested and we would just recommend that they provide two ways of calculating the measure, one adjusted, one unadjusted.

We can go on then to recommend if for accountability applications, those two, you know, both be used in some way and there's fairly obvious ways that they could both be used.

Monica Bharel: It's Monica Bharel here. I just thought, you know, I came in saying that we should not – in my mind thinking we should not present both but as, you know, all are speaking, I can see it compromised. I just need something clarified for me. When we spoke about adjusted versus non-adjusted, can you convince me the difference between clinical and SES and let me just elaborate for a second, so in our two-day meeting, we had very good detailed conversations about how we would like SES factors to be seen in the same way clinical are.

So for example, if we're looking at someone who has a wound on their leg, everybody accepts now that if they have diabetes, that changes the way you look at the healing of that wound outside of any quality measure. And what I am hoping we can do is say that the fact that the person is homeless also changes the way you look at that wound on the leg and healing. And to me, they're equal. Same thing with something like pneumonia and whether they're an asthmatic or not, we clearly understand that as an issue but whether they don't have air-conditioning because they're too poor to pay for their electricity is not currently looked in.

And so I'm having just a little bit of trouble saying the adjusted versus non-adjusted will still take the clinical issues into our account but not the SES side. I'd love some clarification on that to help me understand that better.

Steven Lipstein: This is Steven Lipstein. I'm trying to – I want to get a little bit above the technical again just for a minute because I'm trying to think of what I would do as a large scale provider of services, what I would do with both an unadjusted and an adjusted rate. I can see what maybe purchasers may do with two different measures but as a provider, how would that help me modify human behavior?

And so I have a different construct on this and I'm not opposed to, by the way, doing non-adjusted measures but if every provider, hospital, home care, agency provider could apply a standard methodology that was developed somehow by our – probably by our government and develop an encounter weighted poverty rate based on not where the hospitals are – or providers are located but where the patients reside and the circumstances that they really live in, so that every provider had an encounter weighted poverty rate that they could calculate and understand. And then whether or not you apply that poverty rate to risk adjustment methodologies would be based on whether or not that particular measure lends itself to sociodemographic risk adjustments.

So what – the reason I throw this out as an idea to address this issue is I'm going to give you two examples. We have Christian Hospital in North County which serves a very disadvantaged community, and Missouri Baptist in West

County that serves a very affluent community. They have very different discharge weighted poverty rates.

So what I would like to be able to do instead of comparing Christian to Missouri Baptist on any outcomes measure, what I'd be better sort of to do is say, "OK, where are the hospitals that have the same discharge weighted poverty rate as Christian and compare Christian with those providers to see if there are things that Christian should be doing because they may likely need to spend more money on transitions of care, or more money on discharge planning, or more money on patient education, or more money on home based services in order to affect the same outcomes that other providers with the same discharge weighted poverty rate had been able to accomplished.

But at least what we will do is stop comparing hospitals who serve affluent communities with hospitals who serve impoverish communities as if they're on a level playing field or a poverty-blind basis.

So I think what I would want to see to address the concerns of masking disparities and quality problems is to begin to compare providers and the comparisons are important, who had equal poverty burdens. And I'm not sure just providing me with an adjusted and an unadjusted measure helps me to do that.

Marshall Chin: So this is Marshal. I think that like, in some ways like, one of the reasons why we ended up with the current state of recommendations was that we saw there was this challenge with separating a pure management function which has been – NQF's traditional the role with the application of those measures but wanting to make that bridge.

Then also, there was the nuance in discussion and I think, you know, both of our committed self as well as for the general public end users, we want to have sort of a simple message. But I think in some ways, like the feedback and discussions we've had so far shows that there may be a way to have sort of a simple message combined with the more nuance discussion of how (this might) be used.

So I guess that we come to like to three comments here, one is – I do support the, you know, Larry's suggestion about reporting both the adjusted and non-adjusted. So sort of like the light being the sanitizer effect that just having those numbers out there, it becomes starved that there are problems that need to be addressed and how they're used. So that sanitizing function would go a long way.

A second comment would be that I do think so that what Larry is saying is obvious to people – may not be as obvious to some people and so that – having a discussing about different possibilities for use.

And so Larry is mentioning things like well, you know, the one – and that you could just use the adjusted measure for accountability purposes or you could have some type of combination with the adjusted and unadjusted, or you could reward for (abstinence) threshold as well as improvement and all.

You know, I think all that discussion probably needs to be in there as well as how – the mix of – different ways of using and weighing these factors might need to change over time. So example, Robert Wood Johnson Foundation's comments, they talk about how that over time they would hope that some of the social determinants would become under the preview of a health organization or ACO, or whoever entity. And so the evolution over time of what is considered to be within measures, or plans, or organizations purview could very well change over time.

The third thing is (playing possibly) for what Steve just said, you know, the second bullet recommendation one about stratification – it probably just makes sense to have it instead of a separate recommendation. But Steve has made me think about two ways that we might be able sort of to build upon it.

You know, it's something we had most stratified measures like your typically – well, you know, stratified measures by ways of ethnicity or socioeconomic status and those type of measures. You know, think again, those are really important ways to do things to basically instead of sanitize by showing the light.

But I think that Steve's comment also made me think about – we had an extensive discussion over some of the calls and meetings about – than beyond sort of just the reporting by strata, there are other ways that you can potentially use that for accountability your application for payment.

So I remember for example, people talked about, you could have basically, (inaudible) like, all the (statement) hospitals one been, but then it re-stratify outcomes so that you (are not comparing) apples to apples but then you're stratifying – you're suggesting for – give us the case mix, you know, both clinical and socioeconomic status. So that's the way that this combines stratification as well as in the risk adjustment.

I think those three different elements then, you know, the plan start, you know, unadjusted is number one, number two, discussion about different methods then applying measures for accountability purchase and payment. And third, stratification and potentially risk adjustment within stratification.

Those three I think reflect the – more the nuance of discussions you've had and do it in a transparent way that still gives I guess the users then flexibility and how they actually use these measures, but makes explicit some of the pros and cons we thought of and some of (our) thinking about, you know, it needs to be addressed because I think one thing that's driving this is that currently, CMS, other payers largely don't think about this in terms of complexity of reimbursed form. Well they think about it but they don't implement it. And so I think this will go a long way towards – just moving us further along in terms of action.

Nancy Garrett: And this is Nancy. I just wanted to …

Nancy Sugg: This is Nancy Sugg.

Nancy Garrett: Oh, sorry. Go ahead Nancy.

Nancy Sugg: I just wanted to wrap back around to what Monica said earlier because this is an area I'm having a big struggle with also. And I will say, I'm one of the people that started out with no, there has to be just one number and I'm certainly willing to compromise if it moves us along because I feel like having

socioeconomic as addressed is so important that, you know, I understand compromises need to be made.

But I still go back to what Monica says, you know, there has been a decision that having diabetes changes your outcome in a way that that is beyond just quality of care to physician. And nobody says, "Well why should diabetics expect lesser care? Why are we saying that their bar is lower?"

So that sort of has already accepted and now we're coming and we're saying, you know, we really feel that there is data to say that socioeconomic status affects your outcome beyond what a provider or a clinic can control and it should be adjusted for that. But then we waffle on it, we feel like, well we should sort of put it out there but maybe really not, and that makes me uncomfortable. It's like if we feel like socioeconomic status is a determinant that needs to be addressed in trying to drill down to quality and to compare quality, then we really need to feel that way.

And if I had to choose what I would do is put one (number) out there with everything adjusted that we feel like is appropriate to adjust for the measure. And then allow a very easy route to look at all the unadjusted values from the very beginning, the raw score, the disease severity, and socioeconomic. So then if you wanted to look at disparities, if you wanted to look at it differently, you could but the number out there is adjusted for things that we feel are appropriately adjusted for.

Nancy Garrett: So this is Nancy Garrett and I agree with that comment Nancy. So I also have a concern that if we make a strong recommendation at two different measures, it have to be endorsed. One, adjusted for clinical factors and not socioeconomic and one, adjusted for all of those.

We're setting apart the socioeconomic from that clinic risk factors and we're not taking that stand to say we actually – if there's a conceptual and empirical reason to believe that we should be adjusting for SES, then we should do it.

And so it kind of gets back to the whole – the way that the NQF process works. I mean right now, there's a measure that's endorsed. There's not a process for having two measures endorsed, I mean, you have to submit two

different measure applications. And the endorsed measure is the one that you can pay for performance programs. In Minnesota we've had some experience where our community organization Minnesota community measure that we've created some risk adjusted measures that do adjust for payer status is a very rough indicator of socialized economic status. But that adjusted measure is (buried) on a pay to 150 of a 200-page report and it's not the one that's used for public reporting or for (P for P) programs.

So, if we are – it just feels like we're really moving away from our recommendation if we take that stance so I don't support that idea saying we need to report both. I support the idea of looking at both in the endorsement process and having that strong recommendation not stratification that in order to address disparities we need to stratify and we need to do that process. So, I agree Nancy on what you said.

Female: Can I response to Monica and ask this a question about sort of why would we think about (homelessness) any different than we do diabetes?

Male: Go ahead.

Female: You know, I think this in some ways comes to the cracks of there why we're all in this meeting and then d then (Jeff) isn't holding a panel on should we risk adjust for diabetes or heart failure and I think people know this, I think it's important to go back to it which is that when we think about socioeconomic factors the ways in which they impact outcomes are more complex. So, they impact people through their lifetime and when people arrive at the start of a care episode there may well be impacts that are measurable in the glitch that, you know, people arrive sicker and hopefully we can capture that largely with our (cynical) risk factor but maybe not entirely.

But the other thing is these area is a long history of disparities and care and so although I am confident and when a patient arrives to be care for by Monica or Nancy the issue is not around quality of care. When we look nationally at the outcomes for a patient who are poor or minority, part of what we're seeing when we look at that relationship is mediated by quality.

We know that minority occasions are clustered in poor quality hospitals looking at positive care measures by the paper that she started and that there is, in some cases evidence of discrimination within a hospital around how patients are cared for and sometimes providers are under resource so even though they would or could provide the highest quality these patients are going to under resource facilities and that gets in the peanut piece but nonetheless it means that part of the relationship between SES and outcome is through a pathway that has to do with lower quality and that's what you worried out where as there is now that they synchronize it it's the fact that a diabetic has worst wound healing is not because they are systematically being seen at lower (prior clinic) providers. It is a biologic factor.

So, although, we did not have black and white and a straight clean line, there are some similar issues. I think that we have to acknowledge that there – we wouldn't be struggling with this is if there wasn't more complexity to the issue around SES as a risk factor and that we have to acknowledge that it's not as simple a risk factor because of the history that (very care).

Norbert Goldfield: This is Norbert Goldfield in a positive way I think the comment has already been made with respect to what (Barbara) or (Jackson) said and that's how I respond it all to (Kevin's) question to me which is to say, "You know, I believe that there should be one score in the ideal world it should be one score to start with and over time the institution should be responsible for many acts of the SES disparity. So, I think that's the, you know, that's the way to look at it that over time and initially the, you know, that we should reconsider having both worked together because in fact, you know, the – as Susannah just said, "The line is not black and white," not even black and white with respect to, you know, wound healing for diabetics. You know, available the exist resource of that resource. Thank you.

Mary Barger: And this is Mary Barger, I was one of the people who though maybe the company of two scores, adjusted and unadjusted, would be difficult for the public to understand, but I would agree with it and I think, you know, if once it's done that people will then learn and there'll be education around the two scores and just like we've done with various things that we've done in health

care people learned to what the two scores been. So I would be supportive of doing adjusted and unadjusted.

Karen Pace: Kevin and David this is Karen Pace. I wonder is in line with some of the recent comments that would be the time to talk about some of the discussion we've been having about the methods and the do you it would be useful to do here or do you want to hold on that David?

David Nerenz: I guess those – my first thought is the issues are related and I'm sure they're quite the same or directly linear. I think a lot of it we're talking that with the methods. I think go and see a little more into what I'll call the Susannah line and discussion about the issue whether to do it, whether not to do it and there had just been some good interchange about some of the patterns and data that might see the discussion.

I think want I'm hearing is the issue of whether our recommendations either could explicitly say or at least consistent – be consistent with the idea of public reporting are both adjusted, unadjusted and actually if I can pause it, get the (floor) a second question to you that to Helen and Karen, one of the Nancy comments was that the NQF endorsement would presumably be for just one or the other of those and then it would be or you'd have to actually run both variations through separate endorsement processes and I'm just curious with those most directly familiar with NQF endorsement.

Is that necessarily so or for example could the NQF endorsement ultimately bless an adjusted measure but in some particular public reporting application, the NQF that you endorsed adjusted measure could be reported and also an unadjusted measure which I guess in this scenario is not endorsed. I mean, I wonder if we're just teaming ourselves in this discussion in the last few minutes a little to narrowly into a quarter.

Karen Pace: This is Karen and, you know, we haven't had this particular scenario before but I don't think that it would have to be two different measures to agree, two different endorsements the way I would see this play out if this is the recommendation that when a measure is deemed that it's relevant to include socio-demographic adjustment that when that occurs that there needs to be

stratifications for also computing the – just the clinically adjusted measure and that the two, you know, I think we could say that the endorsement is that the two have to go together.

So I think we could, you know, adjust our framework for endorsement based on these – what – how these recommendations come down and ultimately get approved. So I don't think that has to be limiting factor but, you know, I completely agree and I don't want that be the rate limiting step here that's been just (thought against) but a lot of work that we're really began thinking about how we better integrate the hand off between endorsement and measure section through MAP so I think there are maybe ways that we could consider as creating this inside a single measure and that should be the rate limiting stuff.

Ninez Ponce: This is Ninez. I'm wondering I mean if the problem we're trying to solve right now is to address the concerns that were masking disparities I think it has to be better communicated to the lay audience that you need both. You have to do both to look at if the disparities piece. So I think that's kind of hard for lay audience and maybe some of us to understand but that you have to have both to address on masking disparities.

Sean O'Brien: This is Sean, can I weigh in?

Karen Pace: Yes.

Sean O' Brien: I feel like the recommendation that you have to both is kind of based on accepting this opposition that adjustment for SES factors does indeed mask disparities or have the potential to mask disparities and I don't agree with that myself really. I feel like there is room for clarification.

When we – we're talking of issue of masking disparities, you know, there's a lot of related issues that relate to lowering standards or the, you know, possible locations of incentives and impact to provider behavior and there's really when it come to disparities there's a couple of related issues there's basically disparities can arrive from differences between outcomes within a particular provider whether a different SES category to receiving different care or different center having their needs to meet. Disparity can also arrive

by looking across providers where at certain SES groups tend to cluster within by to it having worst outcomes.

So I think it was kind at the beginning we said let's really just focus physically not on all the related issues but simply on masking disparity. And so when I think about just to asking disparities in terms of disparity it can happen kind of within a provider and across providers where they – where certain SES categories tend to be treated by hospitals that just are not meeting their patient's needs as well.

Just let – Thinking about it literally for the first type of disparities and we've taken it literally does reporting if I, you know, most of reports, most of the measures that we're talking about a reported the unit of a – at the level of provider and we're talking about kind of a single number. So if I know it took readmission rate for heart failure and that's an unadjusted rate.

I've learned absolutely nothing about disparities within (dukes), having or nothing, you know, nothing about disparities nationally whether I now adjust that measure for case next or case next plus additional (FES) factors that we're going to include this case next and that has not really impacted or matched anything related to disparities because it was really nothing – no information about disparities in that single number to begin with. Let alone the issue, you know, separately where we definitely encourage presenting (statistic) results in order to uncover disparities but just see. Literally speaking, the active adjusting and not adjusting at single measure did not mask anything.

And then they did – there's a separate question about well, you know, you may somehow mask the fact that at some particular provider that provider is not meeting its patients needs, and you've masked that by adjusting for SES. And actually when you look systematically across, you know, hundreds of providers you may see that hundreds of providers are failing the meet the needs of their patients and that was somehow you didn't see it because of, of the SES adjustment. And the comment there and this is the comment that was made by (Allan Jaslowski) he's a, you know, the statistician who provided comments. With that, if you're, you know, for the purpose of identifying disparities that are awaited that clearly clustering its patients by provider. Just

a report card that left out a number for each provider are really not the best to – not really the best or most efficient or optimal method we're trying to identify and quantify that type of disparities.

And if you – and it's an important goal to gather information and to identify and highlight it, and quantify it but just to – that's the work we're talking about kind of a different purpose for your performance measures and if you really get the purpose of identifying a disparity, you should be performing analyses that are really ideally suited for that purpose.

And so, given that that's not really the primary purpose we're talking about. I just – I don't see it as a major concern about adjust – about masking disparities because we can always do appropriate other analyses that get out disparities questions a lot better than what we'd be doing just by reporting in August but on adjusted measure or adjusted plus.

(Crosstalk)

Female: So I think in last time, I made up that – I made that same point and I agree that it's very hard to have a one measure that does those things, right? And so, I was sort of against the idea of trying to come up with this magic measure that's able to both measure disparities and account for the (socioecon) effect of course your demographic differences that we should have been adjusting for. That being said, I mentioned at that precludes as reporting those adjusted and unadjusted. And I think part of it and I say that is that well, yes I agree that promoting, reporting adjusted and unadjusted isn't sufficient for really identifying through disparities either across or within groups.

I do think that by reporting both, we also demonstrate some humility in the idea that while we think that this is important to do and that there are certain cases where you might want to adjust for these factors that it's – that someone said it earlier it's not a perfect measure. And so by reporting both it just (steals) a little bit more like for disclosure. And – But I don't – but I agree that by reporting both I don't think we can take the leap to say, oh, no one know about disparities too because I think that requires an entirely difference sort of steps in order to get that to that quantification of the disparity itself.

David Hopkins: All right, this is David Hopkins. That was a great suggestion and just to come back to the – (Larry's) proposal, I think the other thing that hasn't been mentioned is for most of it to who were concerned about the original recommendation and the reporting of the measure only one way is loosing the opportunity to gather more evidence around disparities as they relate to specific measures that are going to be endorsed following these recommendations.

So, I would look upon the research community as an important audience for this dual measure reporting or do a way of reporting the measures. And we will gather much more evidence that way on disparities and how to deal with them. And what actually is going on. So one more reason is support (Larry's) recommendation I think.

Steven Lipstein: David the one thing I wanted to add this is Steve again. Now, especially for the consumer groups I hope it's kind of becoming obvious that in these communities that serve disadvantage patients. There aren't multiple providers so that you're going to be able to compare scores from one provider to the other. This tend to be communities that either have no providers or a sole service, you know, sole source provider.

And so the disparities are really when you compare those that disproportionally served, disadvantaged patients versus those that do not. And so, since today we're basically reporting measures that are not adjusted for sociodemographic factors. If we begin to report those report two ways adjusted and unadjusted, I think most from the provider community would see that as a step forward. But again they're still going to exist at large disparities between communities until we come to grips with that fact that we are still comparing providers who serve affluent communities against providers who serve disadvantaged communities. And that is where a lot of the disparity will continue to exist.

Male: And I think we'll learn a lot more about that David if we do it both ways don't you?

Male: I do, I do. But I do think that the opportunity for performance and improvements still resides in figuring out a way to compare – to find the best practice as observing disadvantaged populations especially in those communities that just don't have a local tax base.

Lawrence Casalino: So this is Larry, you know, based on the recent comments here, I had make two suggestions. One is in societies we have to (shares), but maybe pretty soon we should kind of call the question and once we've done that whenever we do to that I would like to see people discuss the stratification issue similar. And I think arrange to two more words about that I don't want to get into it now I'd rather have the, you know, finish the discussion we're having. But I think that the stratification question is a separate one. I think it ought to be separated and then some people have highlighted both in the memo and today discuss separately maybe.

But I'm not sure we've really exactly made clear how that comes in what we have expect measure developers to do. So I'd like to see that the separate discussion in a substantial one. And maybe we can get to that once we are finished with what we've been talking about.

Ninez Ponce: And this is Ninez, can you hear me?

Male: Well, Dave, can you just give a quick response on the issue of calling the question, I do agree that we've actually still got many issues in front of us and the times gone by. I'm suggesting that which also hopefully is a response that what I hear here are some positive statements but also couple reservations about just the broad kinds of reporting both on adjusted, non-adjusted measures. What I'm wondering is why there as we've done once or twice in the past, once this call is done, Karen, Kevin, and I and Helen we could perhaps jot down maybe two or three different flavors of wording. Perhaps with different intensities or with different nuances about decision reporting both and put it out for the same kind of (structural) ballot we've done in the past. And just see where people are.

My concern is this conference call format is really tough to work with in terms of a call the question thing. I don't even know how we – how we pull the

response other than one at a time by manner take a long. I do agree that just in the progress through this call we're about in a call requesting time. But I think the way I would imagine doing that is kind of a really quick or fine (call) that may allow people to express the use about say two or three alternative reworded statements about this issue.

Male: You know, I would agree Dave. And, you know, I think at this point of, you know, I think we are probably ready to move on unless somebody has a really burning issue that they really feel has not been addressed by previous comments that they want to make.

Female: I think I heard Dionne.

Dionne Jimenez: Yes, this is Dionne sorry. I've been, I'm not sure if you can hear me or not.

Male: Yes, OK.

Dionne Jimenez: But don't know if were going to talk about this during the stratification portion, but I just really wanted to put a big, strong plug in terms of, I think it's very important to have whether or not we have two scores reported but that data is actually available to interested stakeholders, you know, what's the best data both on adjust and adjusted. So if there needs to be different analysis that need to be run back and happen. And I was also curios if David Hopkins wouldn't mind responding whether or not he thinks the approach of having two scores available would help address the concerns of, you know, SES adjust but actually mapping disparity a few things of that would help, kind of the case of making this more acceptable to some of the consumer groups that oppose this.

David Hopkins: Thank you for asking, this is David and as I've already stated the answer is yes I think –and I think it would be a big step forward and the strike for me is the right compromise under the circumstances.

On the question of whether or not to take a vote is, I mean, sometimes there's enough consensus that everybody could agree that adapting in this case, I call it the Casalino proposal would be acceptable, it might not be their preferred solution but acceptable – you can just ask that question, is there anybody who

would not find it acceptable? Is there somebody who wouldn't find it acceptable that we've made a huge stride forward in you wouldn't have to struggle with two or three different ways of framing it and having what votes and what to do then and so forth that's my suggestion.

Lawrence Casalino: Yes, Dave, Lawrenc, here just a quick response, I thought I heard fairly strong reservations objections from both Nancu Sugg and Nancy Garrett. And I wasn't sure even in our discussion of this are we talking about for recommendation that both measures should be, must be, could be, it's – if we have for consensus. So I'm not sure it'd be 100 percent sure what we're having consensus about. I know we're close, I mean I'm hearing what seems like an approach to the landing strip of consensus. But I'm just not sure that it would quite and I'm thinking about our face to face meeting where we eventually got into some very important discussion about the word or versus and took us awhile to wrestle that to the ground.

So I'd like to be able to say yes we got this settled, let's move on, but I feel like we may (steam role) a couple of people if we do that.

Male: Did we have (actually) to restate its proposal and see.

Karen Pace: OK, this is Karen Pace. Because a couple of people also brought up the stratification. So I guess the other option that – and I'll just do a check here that, you know, that of having the sociodemographic it does did score. But having the stratified data available I think that's what Dionne was talking about having the data at large available to whomever would want it. But I guess the other question is whether stratified data by a provider is the other option that I just – I'm not sure if there's anything else to say on this. But I mean, you know, there are other proposals besides this one that anyone wants to bring up I guess is my basic question.

Eugene Nuccio: This is Eugene Nuccio. The question that I have is, is there are an implicit interpretation of a clinically adjusted score versus a clinically in socioeconomic, sociodemographic adjustment. That is if the score with a clinical for an agency is a bad score and compared to another one. And then with the adjustment of the sociodemographic adjustment the score is good.

Are people interpreting that as yes evidence that the sociodemographic is being massed by that second score? I'm just looking at, you know, we're going to have two scores out there and they could be different. What is the meaning of that difference?

Lawrence Casalino: This is Larry, you know couple of points, one is I would start to where you shouldn't just 10 seconds (inaudible), but the word compromise just come up here and there in our discussion so far and I think I just wan to emphasize and this we're leading to a response agent that at least speaking for myself. I don't see this is a compromise on principle. Or I don't see this as something that we would do to make – even though we don't think it's really quite the right thing to do but we would do it to make critics happy. That's not the way I see it. I actually think we made a mistake and I see this is the right thing to do.

And now getting to your point, Gene, I think the unadjusted score if that's all what's out there it's clearly going to be unfair to providers who take care of disadvantaged patients. So I think everybody agrees with that. That I don't think is any of the critics disagree with that. If we put out only the adjusted score it will mask disparities in the sense that and this is I think what the consumer advocates are concerned about, if only the adjusted score goes out there then it is true that a provider organization that is providing worst care for disadvantaged people to continue doing that forever. And nobody would see it and there're being no penalty for it, right? So, to me in putting both scores out there enables one to see the data both ways. You can see how an organization is doing compared to everybody. And then you can see, well OK if it's doing worst, is it doing worst because, you know, because things happen, things – because he take care of more disadvantaged patients.

So I think it's a compromise not on principal but actually trying to balance the conflicting goals that we have. Now, whether they try to, and I think for the measure developers that's probably whether NQF is part of this process once to give any broad advice about, you know, if you got two majors out there how much you used them is I think a separate question. And I think we could also discuss at some point.

Male: Larry can you respond to Sean's comments and I think also – I think it was Monica who raised earlier regarding the points that adjusted and unadjusted may not reflect on the – will not reflect on disparities directly. And that there are better methods to do so, if that is the intent.

Lawrence Casalino: Well, if I understood Sean's point, I would say that and I may not have. I would say that if you have the two numbers and they're substantially different. You know, that is to me is evidence that disparities are important here. And that the fact that one organization is taking care of more, detect an organization has a better score if you're just for SES shows that you think you have a lot of disadvantaged patients. It doesn't show, you know, what the quality of that organization is really. But so you can see I think whether SES is important or not for that organization. I agree that if you want to try to parse out more what's actually is going on. Then starting to try it again and to stratify analogies where you're comparing to others of the same class so to speak maybe very illuminating.

I guess the reason I wanted to keep the stratified on the one hand separate on the other hand give it more discussion than we did. It's because I think there's a lot of problems which are in the report and some of which we discuss. A lot of problems with stratification that to me would preclude being use as the measure so to speak. You know, some stratified measure. It's just too complicated.

And so I agree that throwing things more in stratification can help us learn more about disparities. But I think that I'm, you know, for purpose to reported measures I think the numbers are fine.

Nancy Sugg: This is Nancy Sugg.

Female: Nancy go ahead.

Nancy Sugg: Sorry, so I'm not a statistician, so I will certainly defer to anyone in the team that is. But I disagree that adjusting for socioeconomic will wipe away our ability to look equality. Because if my clinic is doing everything it can to get a diabetic under control. And I know my numbers are not going to be as good as my colleagues in (Bellevue), you know, I will have my number. And if

there's another clinic that has a similar population to mine and will get adjusted and they are not doing everything they can do and their quality is less, their number will be less.

And so I don't understand why we make this assumption that by adjusting then we have wiped out our ability to look equality between providers? And the caveat also I will make with compromise is that, you know, if people feel strongly we need both measure that's fine. I think it makes it complicated for patients to know what to do with these numbers. And if we stratify it will be very complex I think for them to understand that.

But the other caveat is when we get later on to the discussion of should we give sort of regulations of how these measures can and can not be use. We will have to say, yes to that because my fear is exactly what Larry did when he proposed the two scores. He said, well then we take sort of an average between these two score and maybe we use that for paper performance and that's where I cringe. Because it's like, OK, so even though we know socioeconomic status is important we're still going to sort of jeopardize the people that are really trying to provide care.

Lawrence Casalino: So Nancy its Larry I'm very sympathetic with what you are saying. The problem is the kind of elephant in the room on this is that nothing we can do is going to fully wipe out the usual level of inequality that we have in the country, right? And I don't think there's a payment for into that can either.

But I do think that, but let's take three providers. Let's take the provider that doesn't have a disadvantaged population and then two that do. And we only report in the adjusted number for all the three you say. And let's say the – because it's adjusted we see the same number, the same level of performance for the provider that doesn't have disadvantage, right? Then we see a lower number for the other one that has disadvantaged. So, it's true that you say that the provider they think you're a more disadvantaged patients will look better. I'm sorry, the provider that's taking better care of disadvantaged patients will look better on the adjusted score than the other provider that's taking care of disadvantaged patient and in doing as well, but still they'll both look good

compare to the provider that is taking extremely good care of its none disadvantaged patients.

And so they might not be providing as good quality. And they would be able to kind to do that forever without anybody who really been bale to tell if we only report one number. And I think that's the concern that – I mean, that's the main as far as I can tell the main driver of the criticism we received is that that you could do a worse job forever and not be penalized for it. And not even have it be visible. So when I wrote that first paper, you know, in the AMA sponsored paper actually on disparities, you know, in the discussions we had with lots of consumer groups leading up to the paper. The by and large representatives that disadvantaged groups were adamant that they did not want one number. They did not want worst care for them just being risk adjusted away, so to speak.

Karen Pace: So this Karen and I'd like to ask perhaps Sean and others to comment on this. We've had some discussion regarding methods. And I think this is what Sean was trying to say is that this question of – is the national number of being driven – he national difference being driven by for example poor patient's care being clustered in generally lower quality providers so that you're starting with coming up with your expected part of your formula based on, you know, the outcome being driven by those poor quality providers so to speak. So this is the theoretical.

And we've had some discussions with statisticians and other statisticians about methods where you can actually control for that so that you are looking at the relationship of patient characteristic while controlling for provider differences. And I thinks that's part of what Sean was saying that if the risk adjustment method is really adjusting away differences in quality that's not the intent. So I don't know if Sean you can say more about that or others on the call that are involved in methods and methodology can make any comments about that whether that's useful or not.

Sean O'Brien: This is Sean. I kind of feel like that, it was a getting into a little bit of different topic because there's some technical aspects to that discussion. And I think just under basic question of does case mix adjustment pose a problem

with masking. You know, my takeaway message is to know – maybe I'll try to respond. I want to just say one thing just following the discussion in terms of one number versus two numbers. I think it's, you know, incredibly important – I mean I think it's very useful and important to report unadjusted data for a hospital because one way or the other, low outcome at a hospital is – or sorry, at any providers that the hospital as an example can show hospitals that are not meeting the needs of their patients.

And so we need to find hospitals with poor outcomes as important information, but the need to interpret those poor outcomes to relate that equality – is there's been – it's just really not well-defined. There's no – if you see a hospital with poor outcome, you can – you don't – it's either a value judgment that that kind of everybody, regardless of any patient characteristics, should have the same outcomes which is kind of a very ideal way of looking at things. Sorry, I realize I'm not making a lot of sense here but I like to say that I don't see that necessarily evidence of higher or lower quality because I'm not sure we have a well-defined definition of quality. And the way you make comparisons on things that are well-defined as you compare apples to apples.

And so when going back to this point of one measure or two, if I'm going to import two measures, I would consider reporting one that was fully unadjusted because one way or the other, it is important to know about hospitals that aren't meeting their patient's needs. I don't see why you need to adjust that for anything.

And then, last thing I'd say is that all that make – so I'm like I'm saying that our recommendations didn't go far enough. I'm actually not. I feel like it's important to (piece) out where did we go too far, whether there is any overreaching and, you know, for me, I feel like we're – there's – in the report, there's a lot of should statements and then because the should statements don't hit the nail on the head, we have to follow with more should statements. So we should adjust our socioeconomic status that is not perfect so we better do something else, better report at both ways and we should, you know, when appropriate, we should adjust – sorry, I know it wasn't that lack of nuance, but, you know, we basically imply that it's on the developer's burden to show

that they don't need to adjust for socioeconomic status so we put in additional caveats if the factors are not primarily mediated by quality and that distinction may turn out to have some problems.

So, you know, although I may sound like I'm advocating stronger recommendations, bottom-line, I kind of say the recommendations that leave a lot of decision making to the developers.

Monica Bharel: This is Monica Bharel here. I want to just build on a couple of points from what I had said earlier that Nancy Sugg and Nancy Garrett and now Sean are building on. So, I would favor us just for a minute thinking about if we report two, can we report fully unadjusted versus adjusted? Sean and others have, you know, the question that we're looking at here is the theme of does SES adjustment mask disparities. And Sean is saying, from a statistical point of view, no. Larry has said that if we have the two different numbers, what they will tell us is what the burden of care of disadvantaged patients as opposed to quality per se.

So if that is the case, then why not report the fully unadjusted versus the adjusted. Along those lines, I just want to bring one more factor and I know we're not talking about pay for performance here, but the reason – part of the reason that we have 670 comments and this is a hot issue is the cuts we're talking in the end about money. And when we – we're trying to not worsen disparities or mask disparities but the truth of the matter is that pay for performance programs can worsen gaps and disparities when programs that take care of people with – like hospitals that take care people with lower SES, then are getting less money. And what they cut from their programs and the example of this all over the country, they cut things like their food pantries and their witnesses of violence programs and other things that are needed by this population. So, I just want to bring in that point it's really going the other way also can have an impact on disparities.

Susannah Bernheim: This is Sue. I just completely want to support Monica's last point.

So I wonder – I don't know if Karen and Kevin and David can give us some help again on this. I mean it's come up a bunch of times. This is Susannah,

sorry. Because I think there are – I think mostly we all have concerns about fewer resources going to institutions that if anything may need more resources in order to meet the needs of their patients. And it's very hard to separate these things and yet, you know, programs can use measures in different ways, right?

So there are programs that CMS has right now, their case that – I'm blanking on this. It comes out (inaudible) in my mind which explicitly to these places that do badly on readmission measures and gives them fund to try to improve that, right? So, in that case, the, you know, the worse you're doing the more likely you are to get funds. So you can put a measure into a program in lots of different ways. The number of people on this group has suggested using improvement course and so how the measure is designed and how the program is designed is not the same thing but there is this concern that anything we do is going to have implications for those payment programs.

But it would help if you can guide us about what we should do about that in this context because I don't want us to make recommendations that are really meant to be policy recommendations and use the measures as a means to do that. But I also think that there's a lot of examples that people are reflecting on where the way the measures get used now could – is concerning.

David Nerenz: Yes, Dave here. That's a wonderful point, I mean, to quick response and then we'll see what others say as well. I think we all understand and agree that out there in the world there is a very close connection between measurements just from a technical activity point of view in money. And the reason there are 670 comments is money is at play. No question.

That said, I think in general, we're on more solid ground when we try as much as we can to focus our discussion, recommendations on measurement and if we have to make some observations about that links then to money and what happens in terms of fewer resources and the city providers. I think in the reporting, its current form, we make those kinds of observations when we are trying to make the case that the lack of adjustment matters can indeed cause harm.

And I think in terms of general placement, that's still not a bad place but if I'm thinking of your line of thought her, Susannah, correctly. To the extent we can keep our discussion and our recommendations focused on measurement and on the validity of measurement on the informative value and the validity of comparison across plans and providers given different approaches to measurement. I think that's one we're closest to the charge we were given, we are closest to NQF's essential role and their scheme of all the other actors out there.

It's certainly fair as we talk about context, or as we talk about implications to get into issues of money, resources, what not. But I don't think we're on our center of the target or on solid ground. If for example we start talking about programs that give additional resources to intercity hospitals for example, that's really what we are asked to about here. Yes.

Lawrence Casalino: David, this is Larry. I mean, I broadly agree with what you're saying and I don't think, you know, we don't have a choice to figure out to get more resources to intercity hospitals, right? But, the comments that were just made, I mean, we really do highlight in a very precise way. The fact that for this – in this case, you know, distinguishing the measure from thinking about how it's going to be used, I think is impossible.

So – I mean, let's just take pay for performance very specifically. And this is why I think this is a compromise on principle, not a political compromise and that what I suggest is, if we were to use – if we were to recommend just an adjusted measure as we did and that were to be used for pay for performance as it will be, right? Then we are protecting providers that take care of a lot of disadvantaged patients which we really want to do, right? We don't want them to be unfairly heard by a pay for performance.

So – but on the other hand, if we – if the measure is just an unadjusted measure, providers that take care of a lot of poor patients can forever take worse care of them and not be penalized financially and, to me, to say that that's OK would be just kind of a political decision that while we just want to get more resources to those providers any way we can, which we might all agree with, but I think that does go beyond NQF. To me – I mean, again, I

think the compromise is between recognizing that whatever number gets put out there or numbers is going to be used for pay for performance, if it's just – if it's one adjusted number, it's going to raise concerns and concerns among consumers that – and among disadvantaged groups that you'll be able to take worst care with me forever and nobody will never know it and you will never get penalized for it, right?

(Crosstalk)

Nancy Garrett: So this is Nancy Garrett.

Steven Lipstein: Larry, this is Steve.

Lawrence Casalino: If I just – Steven if I could just finish one more (inaudible), I'm sorry.

Steven Lipstein: Sure.

Lawrence Casalino: I think someone is said earlier and I don't think you got enough attention that I wish I knew who it was. That we don't know that much about putting out the two numbers yet and that could be a recent argue against doing it. What the individual said, I think it's true is there will be a lot of energy into I think how the two numbers can best for use, how can we best understand them that I think we're be very productive and we'll lead to some progress on the issues that we're talking about today.

Steven Lipstein: So, Larry, I was going to say was the measure itself and how you produce use is one aspect of this. The second is that when the measure is applied in (paper) performance program or accountability programs, what gets rewarded or punished is the variability of the measures among providers.

So what we're punishing is variation, not the score itself. And so, if the variation is explained by clinical factors or non-clinical factors, sociodemographic factors or the lack they're of, it becomes a really, really important dimension of this.

So, what if, you know, what we recommend is that in the measure development process, the measure developers show their measures both

adjusted and unadjusted and as part of the NQF endorsement process, then there's a decision making apparatus that decides which of those produces the more relevant comparison of variation in terms of trying to encourage performance improvement.

So, I mean, I think that producing both adjusted and unadjusted scores to inform weather paper performance should be poverty blind or poverty informed is an important thing to do. But I don't think we should go to the point of saying that the variation is punishable or not.

Lawrence Casalino: Steve that maybe the case but I think asking the measure developer to develop the two measures and then asking the committee that they're doing the measure to pick one of the two. To me that just kicking the can down the road, you know, it isn't different recommendation, and it's letting each committee then that reviews measures makes its own decision each time.

I don't think that will work very well. I mean, I do think we have a decision about should we recommend that one or two is included by the developer and then improved or not approved by NQF.

Steven Lipstein: Just to clarify Larry. Are you suggesting that in addition to the adjustment measure that there will be transparency and the unadjusted on people have raised, well, maybe that should also include measures that are unadjusted for clinical status or even just pure raw measures.

So, can you clarify exactly what the proposal is that you're making Larry?

Female: Yes.

Steven Lipstein: Is it just two measures or is it making the data behind this more transparent and available.

Lawrence Casalino: I'm proposing that measure – that if – maybe very, very specific. If the measure depends in some reasonably meaningful ways on the patient having to do something that the measure should be – that SES is been a factor. And when SES is a factor, the measure developer should tell NQF how the

measure will be calculated. The score of the measure will be calculated in an adjusted and unadjusted way.

And when I say unadjusted, I mean we can debate this but I would favor that the unadjusted means – yes, adjusted for clinical factors but not adjusted for SES and the adjusted measure adds in the adjusted for SES. Now I recognize there are some technical issues with (inaudible) and so on and so forth. Not sure that should hang us up right now.

So, unadjusted is clinically adjusted but as the SES adjusted. Adjusted is SES adjusted. If SES is relevant, they going to tell us how both would be calculated. And then what the further issue then which we may or we may not want to address and possibly not how the world would these two measure, but I think there will be a lot of productive debate about – how the two measures would be use.

And I think people are not that dumb that they can't understand two numbers. I don't think.

Nancy Garrett: So this is Nancy Garrett. I just want to add couple of the last comments. I think – when (inaudible) about a lot of the discussion we're really accepting the assumption that we're asking disparities by doing the SES risk adjustment in the cases where there's a conceptual reason to do so in empirical events.

And, you know, we've taught a lot in this committee about balance we're doing here between the risk of (harm), related providers being penalized for taking care of this damage populations versus the risk of potentially in asking disparities.

And so, I would favor an ultimate proposal which is to really strengthen the disparities and the stratification recommendation of recommendation 1. I think we should pull it out as a separate recommendation and really emphasize that if you want to analyze disparities that's absolutely something that we need to do giving a single number is not the way to do that. The way to do that is to stratify to understand the groups we're looking at and what are the differences. And then dig in to why those (inaudible) that they're happening and trying to address the disparities.

So, I would favor the kind of what's in the (inaudible) here about pulling that disparity recommendation out and strengthening it. I am concerned the (pain) of how do this if we're really going to recommend two numbers.

You know, we've got 56 out of 68 comments (inaudible) support of recommendation 1 with no changes. And so we need to balance that. I mean we're getting a lot of support for the way we put this forward.

Karen Pace: This is Karen. I know we promise to break, so maybe – should we – I think we should probably take a quick break and let people get up and move around and take care of things if they need to.

So, why don't – it's almost 1:15. Why don't we reconvene at 1:25? And we'll leave the line open but this one – people could put their phones on mute. And we'll reconvene at 1:25

Male: Good. Thanks Karen.

Karen Pace: OK everyone, this is Karen. I think we'll try to reconvene David and Kevin online.

David Nerenz: I'm here.

Karen Pace: OK. Do you want to start to talk in terms of how we should proceed? We have the good discussion about the two-number issue but what are your thoughts about moving forward.

David Nerenz: Dave here. First of all, again thanks everyone. It's a difficult process to not be able to see each other and read smiles and frowns and (inaudible) and hands in the air and that sort of thing. And, you know, ask before people than productive, people have been (courteous), people been respectful, lots of good ideas and I think this is been really good.

We've diverted this. We might expect a little bit from the plan sequence of topics. But I think its extremely good and useful and I think we've done fine because basically what we've had in front of us I think is the significant topic

of discussion that may bring us closer to some common ground between the original high level on consensus in the panel and all the positive comments that we received all that in a one hand and the concerns by purchases on the other hand.

If we can find a way or two in which we can speak to some of the concerns by the purchasers without sacrificing principle without breaking up the consensus that we've had in the panel. I think it goes without saying that's a very good thing increases the chances of our recommendation to actually (inaudible) that the NQF board level.

So the amount of time we spent on – I'll just call the two-number topic, I think it's been extremely good. I think we still have perhaps a bit of words and I think – and checking with everyone to do on that point.

My suggestion I think would be to try to use the remainder of our time on two or three other significant issues that Karen had hit in front of us in terms of the memo on the slides will obviously have to work to those a bit more quickly that we might have otherwise – but I'll say I've been very, very please with your all done so far today.

Kevin Fiscella: Dave this is Kevin, I would echo Dave's comment. I think that this is a really important point and I think it was (less) the time that we spent in questioning on everybody (feels on it).

Male: Kevin could I suggest the next big question that maybe we could address and maybe a little more briefly. But is the question of when to apply these principles, that's going to raise a number of time – and it's on the context of a measure developer scratching their head because they didn't originally come to approach whatever measure is that they're working on with the thought that sociodemographic factor should be incorporated in the measurement.

It occurred to be as I've read and reread some of the stuff we've been through in this journey that when we first started all, I think that some people were concerned about those measure developers who have been very eager to include SES in their measures and NQF is prevented them for doing that.

So recommendation 1 as a I read and reread it finally hit me that it was probably more addressed to that situation enough that around to a developer that if not inherently thinking about incorporating SES.

Under what circumstances are we saying to that developer, you've got to do this, whatever this is, whether it's a two measures, two variations, solution or whatever. But if there are some ways, we can be more specific about that.

David Nerenz: Yes, Dave here. Again, quick response. Yes, indeed. And I think our challenge maybe to try to leave together a couple of very closely related lines of phone call and e-mail exchange that's been going on in the last few days. And let me see if I can (tee) this up open it and then – Karen, I'm hoping you can do the glue that hopes us together here in terms either what appears on the PowerPoint set and I'll try to do the best I can.

There has been some discussion that included in phone call yesterday afternoon with the statisticians in our panel as well as Alan Zaslavsky who I think most of you know who spent a lot of disparities work based at Harvard.

Where this got started is during the open commentary and I had an e-mail exchange with Alan, as you know from his comments, very supportive of our report. But he began thinking about some analytic methods that might speak to this question of at the front end when is it appropriate to do adjustment and then whether is appropriate or not. And as a non-statistician, I'll say it fundamentally had to do with some issues of what sort of patterns to do you see within provider variation or disparity related to a factor like poverty and how does that related between provider issues.

And there was a lot of e-mail exchange. There was a phone call yesterday and I think I made at some point here turn to Sean to talk to that little more. And I think that –and Jean as well. So I think this a good time to get into that.

I think also then we had to open the door to Susannah who's been thinking about these issues and also hit – send some things last couple of days. I think related to this point and it just also I think speaks to the point that she made in the opening of the call about framing our discussion in terms of when or when to adjust rather than more simply yes and no.

The art here and trick is obviously how do (leave) two things together if it is indeed the time to do that. For those who are involved in either of those threats, open to suggestion.

Karen Pace: This is Karen. I think that – but sounds good, Dave. I do think that the thing you brought up about the discussions with the statisticians is relevant because it directly speaks to the concern about risk adjustment, adjusting away the quality differences or the fact that poor people are just generally getting poor quality and care.

And I too will ask if Sean might talk about that in a little more detail, so that we have some understanding of that or discussions.

Sean O'Brien: Sure. This is Sean. I'm going to try to relate some of the comments from Alan the statistician, Alan Zaslavsky and our subject of discussion. But as (back on) I'll just say two words about how I think some of these issues. The concern about masking disparities I think some of the comments reflected not necessarily – I guess I distinguished kind of concerns about scientific issues and clause statistical methodology that might have an intended statistical effects, so you want to measure something but you end up addressing away – adjusting away the thing that you want to measure and you apply these counter reproductive critical adjustments versus just a more general type of objection that there's something that hits us in the face about applying adjustment that kind of imply the double standard (admit) and seems to have kind of some ethical uncomfortableness that we may apply adjustment that on the surface appear to value different patients differently.

But really then, you know, I think the question that Alan was addressing was from a statistical perspective is there some flaw in adjustment, would be – if we start with a goal it basically measuring provider differences would we adjust the way those differences by accident by doing case-mix adjustment.

First point he made that disparities can arise from two different mechanism and it's worth while distinguishing them. And you can think about within unit, like a units being a hospital physician or provider. You can think about within unit disparities whereby members of disadvantage groups receive worst

care or have worst outcomes in other patients within that same unit. And then between unit disparities whereby member of disadvantage groups are more likely to receive care from units who are overall lower quality.

And the point he made in his initial comments and there's more discussion about this. Basically, you don't end up adjusting away between the quality differences you want to measure because (as long) within unit effects that are estimated and control for when you do case-mix adjustment.

So you basically by ensuring that you kind of always comparing apples to apples that is – will allow you to control for differences within providers but it's really between providers (inaudible) estimating and they won't adjust that away.

Then I think when (care) and Alan had a further discussion, he pointed out away that case-mix adjustment is like often implemented. It's not so obvious that it's always adjust – the risk models are always explicitly adjusting away within provider effects in order to make a parent between provider effects.

And so the statement that, you know, that we're OK because we're always controlling within unit effects is there are some difficulty with that. And it turns out that, you know, basically some statistical adjustment may have unintended effects.

I guess the message would be that – it's not a property of case-mix adjustment in general. There's certain things that, you know, (inaudible) pitfalls and limitations and the type of analysis really need to be match up with the objectives and handled appropriately. But there are – they're definitely statistical methods that do not lead us by.

Just to recap, I mean the issue we're talking about is the same when we're already talking about where patients may tend to cluster in lower quality providers and – just the bottom line, you know, cause and start again. But the bottom line is that there are appropriate physical methods that can piece out that within unit effects and make apparent between provider quality differences that we want to estimate, but not all of the analysis that are done

for performance measures are actually doing it – what we might on the surface expect they're doing.

David Nerenz: Dave here. And I guess trying bridging us a couple of things. I think even though this discussion for awhile maybe probably technical. I think we're touching on dynamics that Nancy Sugg's capture quite eloquently before the break talking about if you're serving disadvantage patients and the numbers don't look good, it could be, A, because you're not doing a good enough job but also just as rationally you could be – you're doing a wonderful fabulous job. But some aspects with disadvantaged are still affecting the outcome and it's outside of your control. And as we slide back to a technical side, the issues can that distinction in causes be reflected in some patterns between within unit and between unit variations. And there seems to be some hint that perhaps they can and that we didn't really have that in our discussion earlier.

And I – there are a couple published papers on this point that sort of illustrate the pattering. So even though this could get quite detailed her for a while, I just want to reassure that non-statistician folks that we're not going to venture too far from objective reality that most people can experience.

And one other thing is that we're going to use the term provider (friends) in here but for (friends) from (AHIP) who maybe following us what we talked about here as it relates to plan as well as providers of any type.

Lawrence Casalino: David, this is Larry Casalino again. I mean, the question is – I think the question you started off with was when should something be done about risk adjustment, is that correct?

David Nerenz: Yes.

Lawrence Casalino: So to me the guidance would go – the developers would go like this. Tell us whether you think the measure you're proposing, whether performance on the measure will be affected by things the patients do or not, and tell us this based on evidence if you have it and on logic, right? So logic tells us that CLABSI does not depend on patient behavior and should – therefore,

plausibly not be subject to SES risk adjustment, whereas other things, pretty obvious logic and/or by evidence might seem to be.

So tell us – you know, tell us whether you think that it needs to be adjusted or not. Give us your arguments? And then if you – you know, if you think it doesn't it – it ought to be adjusted, tell us how you're going to calculate each of these two measures? And that's it pretty much.

David Nerenz: Yes. And actually, I was – Dave back here again. I was taken by your very clear phrasing in this, you know, if there something a patients do that matters on the outcome. I'm not sure we have that so clearly captured before.

I'm thinking as we go along here, whether it's going to be possible to bridge that clear and simple concept to some of these other more technical things that have been flying around last couple days about the within and between and at a gamble here, but hopefully, going to a good place.

Susannah, does any of this tie into those thoughts that you had on this sort of basic question we have of how do you deal with the when and when not and what sort of guidance we can provide?

Susannah Bernheim: Sure. So, I think one thing that I want to say just for clarity is that I think that when we talk about the idea of mapping disparities, there's been some – we have (inaudible) in the in-person meeting as well. It's not the most useful phrase because I think – actually we don't know what we mean by that.

But I just want to explain conceptually what I think is people's greatest concern about what the effect of risk adjustment for SES's which is that the way many measures are constructed, you have an expected rate to let you – what you put into your risk adjustments creates the expected and you aggregate the expected for all your patients and you compare those in some version of what actually happens with all your patients, right?

And so when you set the expected, whatever your risk adjust for is setting what your patients are compared to. And so the concern is that if you set that expected with SES included, and if in general, SES patients are doing worse, you then can have a provider who objectively have worst outcome whether or

not that's because of worst quality, and a provider who objectively has better outcome who are judged to be the same.

So they are both considered to be doing as expected or average or considered to be doing good. And one of them – if the patients are actually experiencing much better outcome than the other ones, these patients are actually experiencing much worse outcomes because that provider has the one with the worst outcome had more low SES patients.

And what I think – originally, when the concept is masking disparities that came up, it was that disparity and outcomes that becomes invisible that this provider with more low income patients, they are patients have worst outcome which is a little bit different than masking quality which I think …

Male: Yes.

Susannah Bernheim: … is also a concern. But I just wanted – I think we mixed or we masking quality or hiding quality differences with how we're hiding "disparity." And there's no question whether you think it's a good thing or bad thing that if you set a different expected for your low SES patients, a facility can look just as good as another even though their patient in fact our experiencing worse outcomes just because they have lower SES patients.

That doesn't tell you whether that's the right thing to do or not. But I just wanted to reminds people that I think the – I think that the root of some of the concern about masking disparities was what you don't see is that this hospital is ain't called just as good even though its patient are doing worst because they have low SES patients. And that to some people and it is uncomfortable.

The quality piece that I don't know quite, I'm trying very hard to link this to the pieces that Sean was saying because I don't want to divert us. I think the quality piece is whether or not those differences and outcome and then sort of thing the same thing again but I think it's just repeating are primarily due to inherent differences in patients that can't be influence by providers or inherent differences and quality. And it seems to me that if the differences are predominantly due to differences in quality. Then, that is going to happen

probably at both or within and between unit levels and you will rule at least some, if not all of that.

I'm going to pause because like I'm worried that I'm not answering the question you want me to address.

David Nerenz: No, no. Actually, this – the very last thing you said is sort of telling in fact that, you know, we've had some concerns and some written comments about this phrase if not primarily mediated by quality. And we have been challenged to either remove it or to explain what it means.

And what I picked up myself in a couple of this recent days discussion is, and on I'll slide back, I hope right directly to your phrasing but something – if you in fact, actually is mediated by quality, if you're looking at a dataset in which you have both within and provide – within in between years effects, you should see both. However, if you only see within, you see no between that suggested the issue is not primarily mediated by quality.

You know, I've just oversimplified this by a ton and Sean is probably (cringing), but I guess that's why you see some of these things coming together that there actually are some patterns in data that we can describe without getting way too technical in any revise report, that actually make this verbal concept a little more concrete and actually might provide some guidance to developers and future review committees.

Susannah Bernheim: And maybe it would help if I could ask Sean just quickly, if Sean (must) take black and white patients, we knew in an extreme case with a particular outcome that the entire difference in outcomes between black patients and white patients was the inadequate payer for the black patient rather than going to worse quality providers where they are being differentially (inaudible) in private. We knew that those differences in outcomes were due to differences and quality. What is the effect of risk adjustment for race?

Sean O'Brien: Well, to the – in a very ideal hypothetical like that, I think you could get away without adjusting for race. But I would probably make the point that if you adjust for race, you're still going to be picked up. You're going to be able to still identify providers that are doing better or worse at caring for their patients

and that, you know, the outcome is kind of – are agnostic to the mechanisms that lead to good or poor outcomes.

So you can, you know, it's everything is relative and you're basically going to pick out the providers who are doing better, you know, better than typical on black patients are better than typical on white patients and it – and may not be typical, may not be the ideal standard that you would like to compare to. You'd like to compare to maybe some standard of every, you know, how's the outcomes of income patients or, you know, of the least advantaged groups, but, you know, they're kind of currently set up to compare outcomes to, you know, across providers, and for that purpose I don't think and adjustment for race in your examples would prevent the ability to identify better worst providers.

And I would say also the answer a little bit depends on, you know, whether you're looking at doing a direct or indirect standardization of I think …

Susannah Bernheim: Right.

Sean O'Brien: … some of the …

Susannah Bernheim: But if you saying that typical is for quality, I'm not worried so much about whether I can have to doing better on for quality or worse on for quality. I am going to set typical at, you know, exact which is in my – I do like scenario is worse quality, right? And I think that – I guess I'm trying to really make clear with the risk is even that we know in this country that some of the worst outcomes are based on a poor care and it obviously differs by the SES factor and the outcome you're looking at.

Steven Lipstein: This is Steve. Just to take this a little bit out of the theoretical back to the practical for a moment and using Susannah's example. If you get different outcomes between the black patient and the white patient, and you're trying to describe that to a differential in quality, in the real world, you can infuse a whole lot of resources to try and offset that that differential outcome.

But then if you produce the same outcome for the black patient as the white patient but you spend twice as much to produce that outcome for the black

patient, you get – there's another measure, it's a value based measure which says that the quality of the outcome divided by the cost of the outcome equals value. You're still going to punish the institution that spends twice as much to produce a better outcome.

So it isn't just a quality differential and I think to have the conversation about quality independent of the resources available to produce outcomes is to say, let's pretend the world is the way the world really isn't. And so, I do think that a scribed difference in outcome, just the differences in quality is a really hard thing to do.

Tia Sawhney: Well – and this is Tia. And keep in mind, quality is a broader concept than just one particular healthy vent. So the outcome of a hospital stay maybe a very dependent on the quality of care that the patient has received for years, an example, a birth outcome. You know, that's going to dependent on the prenatal period and perhaps a mother's entire health history. So to say that our hospital has for birth outcomes and it's raise related doesn't – it could be related to the quality of mom's care but not necessarily what happened in the walls of a hospital.

Male: Right.

Lawrence Casalino: Hi. This is Larry. Operationally, what discussion are we having and why are we having and I'm not exactly sure that I understand.

David Nerenz: Yes, and Dave here. I realize that sort of start one thing is that I think the – the thing I'm trying to pull together here is back when we were working on the wording of the draft report, we introduced this phrase not primarily mediated by quality, and I think Susannah suggested it, although I think I've certainly favored it a number in agenda report.

And then in many of the comments, we've had suggestions back from all kinds of different people that I'm not sure what it means. I'm not sure can be operationalize. I'm not sure how that provides guidance to developers. What should we do with it?

So, I think some of this discussion is about that and the reason that the various technical discussions including the within and between came up and I think are relevant, is I think that it represents, although expressed in technical terms sort of an answer to that challenge that there actually are some patterns and data that can stick to this question of what that phrase means.

And then I think while we're on the same topic that this – the analytic methods that look carefully at within versus between factors also speak to this separately stated concern about masking disparities and that's where, for example, that Susannah questioned to Sean a few minutes ago is in a certain scenario, is there actually a method by which you could adjust for race that would not completely mask quality disparities related to race.

So, unfortunately we've got a few things flying around here at the same time and we've got a technical overlay but somehow these seems like the time in our four-hour (slot) to try to work through with these because I do think that we've got a couple of positive entities in front of us to speak to. Concerns has been raised either about this not primarily related to quality question and about the masking question.

Lawrence Casalino: So David, yes – it's Larry again. That is very – thanks, that helped me a lot. You know, just one thought and I don't have – I don't think I made – I wouldn't presume to take a position on this but just to throw it out for people who can think about it better than I can, I mean, the phrase, you know, that not primarily mediated by quality, it is, you know, it really opens a can of worms as we've discovered.

And I wonder if the concepts that, you know, going back to that, just simple concept just logically, is this something where the patient's behavior is likely to affect the result or not. IT could be a substitute for that phrase in effect, and again I don't mean to take a position on this because I think I would be foolish to do so but could the one which to me seems pretty straight forward. I wonder how much disagree would there ever will really be on whether patients doing something or not doing something affects a result. Could that substitute for this very thorny issue of, you know, not primarily mediated by quality?

David Nerenz: Yes. And I think that's a fair question. I think we're just exploring approaches that the group can think of to address some of these challenges and criticisms that are made.

Female: Could we also give examples that would illustrate our thinking? Like we have talked about the difference between central line infections versus readmissions and the fact that the role of the patient – and the impact to the role of the patient involved.

Male: You know, I think some of that's in there but it's just not highlighted very much.

Susannah Bernheim: And I'll say on this role that – this is Susannah. You know, when you highlight a case where the patient has very little role like (class C), it makes a lot of sense that that might be an example where you would not want to risk adjust for SES. But I would say when you do the reverse, that's a very problematic task, to say we as providers – if the outcome depends on a patient doing anything then, we're less responsible when there's a lot of data that sort of what we do at discharge from the hospital when we're giving instructions, when we are talking about medications, which medications we choose, has a tremendous impact on what patients do.

So I wouldn't say if there's any role of the patient then you got to risk adjust for SES. I think that's going to be a problem.

Alyce Adams: This is Alyce, I agree with that – sorry.

Female:: Go ahead, Alyce.

Alyce Adams: I would just say I fully agree with that. I mean I think that unfortunately it's not less thorny than the quality of care comment that we made before. And as I recall the reason why we stated it that way was to say that we don't want to mess quality differences and let's just say upfront. I understand that it's hard to say there's a – OK then what is that and how you measure it, and I get that but I'm very weary of this idea of saying, "Well if patient's behavior is really

sort of the determining factor here.", because of the reason that Susannah just sort of outlaid.

Male: And this is (Inaudible). If I could just go back to Steve's comment on, you know, it's twice as hard with some patient's as others. I mean I agree that with any patient, if the provider puts in insufficient energy, the results will be improved. But to say that, you know, if you take someone who speaks another language and has a third grade education, and maybe different cultural beliefs and no transportation, and on and on and on, and you say, "Well the providers should be able to overcome all those problems and get just as good as score as they would if they were taking care of someone like the people on the phone here." Yes that's true, in theory you could, but in practice it's pretty unfair I think to the providers.

You have to take care of the patients that – or choose to take care of the patients that needs so much more help. And that again why I think the concept of it does what does the patient real matters is important. It's not to try to excuse the providers but it's to try to say that it does matter to the results on average, what kind, you know, how hard it is to take care of the patient.

It's a little like saying, you know, there are poor people who grew up in the barrio or in the ghetto, and get into Harvard. So therefore, we don't need to actually worry of that giving any extra help to those disadvantaged people because look, in fact, there were comments like that were some health plans are able to do just as well when they take care of disadvantaged patients. Well that's great. But on balance is this, you know, disadvantaged people have a tougher road to hoe and so are the providers who take care of them.

Nancy Garrett: This is Nancy Garrett. You know, in the report, the phrase, "Not mediated by quality." I was one of those people, I felt that was confusing. I don't think we really needed that because I think we very nicely laid out the – we need to be – there needs to be a conceptual link. There needs to be empirical evidence as a link. And I'm not sure exactly what task we're trying to do right now, but I don't know that we're going to be able to define very specifically exactly what that conceptual link is going to look like because – I mean this is – the (science) is early on how to do this well.

And, you know, we have such limited time and we have a report due in a month. I feel like by saying what we said, we allow then future people working on measures to really get into the detail then figure it out. But I don't know that we can neatly give the criteria for what that conceptual link would have to be.

David Nerenz: Dave back again, I'm just sort of kind of monitor the extent we are making forward progress for right now. And I see this is kind of bumpy waters here. I'm wondering, you know, this is largely (inaudible) suggestion to Karen, Kevin and others. The statistical discussion is that kind of complex so I don't think we have to made an effort to get any kind of written summary in front of the people in terms of how it might speak to this question of either the masking disparities or the phrase, not primarily mediate by quality. Maybe that's a challenge that the (304) should take up sort of immediately after this call, next day or so. Put some things out to a whole group that they can think about and react to.

And basically, we would be talking about proposals for things that could ease – be woven into or vice versa and report. Because I realize as we go through to this discussion right now, that either I didn't tee it up all too well or there are couple missing logical links that we need to make sure we're clear to folks.

I think in my mind there's a positive and useful connection between some of the technical issues that were included in e-mails from Helen and others. And some of the challenges were given in the comments that we're obliged to speak to. But I think for the whole group who's not seen all that, we just need to make those connections clear. And then, sort of raise the question of what do we do with them. But I got a feeling we can keep bouncing around here for quite awhile and that make much more with progress.

Karen Pace: This is Karen. That sounds good. And I think, you know, part of it will be just useful in responding to some of the comments, but I think that's fair enough. It is hard to talk about these things when everyone has materials in front of them much less kind of in our conference call environment. So, I think that's a good suggestion for us to move on.

(Rick): This is (Rick). One – I agree with that take that the follow up separately. But I think the (inaudible) discussion with addressing a very specific scenario. Would that had to do is basically the tendency for different SES groups to cluster within providers and the question like, "Wow, could the people are raising these issues about masking disparities be right after all. But that was a very specific scenario, and even taking a way that scenario – and I think there's some assumption that maybe that knowledge of that scenario and concerns about that scenario was it was driving a lot of the, you know, the majority of the concerns about masking disparities that were received in that comments.

And I don't actually think that was what was driving the concern. So I think that wouldn't – we have Susannah for her kind of take on the approach. I mean that's that first consideration she raised was really that measures are calculated usually with the model that calculates an expected value. You know, the expected outcome rate, and that if you estimate that model including SES factors, you're going to, you know, potentially have a lower expectation for different SES subgroups.

And that, you know, may not be desirable for incentivizing improvement, it might be uncomfortable ethically, et cetera. But that same concern is present without respect to any this physical discussion we're having. So I just say that not sure that we need to bring that up anyway.

I think there were some nuggets in the discussion that can definitely be use to inform and better respond recommendations and perhaps even refine recommendations but I don't think it was going to solve the people's – comfortable enough with the masking disparities issue and I think we've instead given, you know, if what I'm saying is right that's that people were not identifying that specific scenario and that not the root of their concern.

And it is more general concern that come, you know, I think (you need to be) more a little (note) more to critical reflection on what does it mean to say that we don't want to accept lower standards and does that resistance – what problem you run in to, and inconsistencies and more discussions of, you know, basically the same point that Steven made and if you're saying that

you're expecting equal outcomes for all patients, is that – you know, what you do and there's a risk take twice as many resources to achieve the same outcomes in group one compared to group two.

And, you know, the basically, you know, my feeling is that the physician implies a lot of value judgments and valid just in (inaudible) that, you know, makes them explicit that the way, you know, is statistically we, you know, we try to get away and rely in a huge value adjustment as a possible is that when there is consensus on value adjustment is not a problematic and everyone on this phone call would agree that for any SES subgroup is take any given patient with the set of characteristics, the brighter that's going to take care to the patient the best is a better providers and the one that gets the better outcomes is better.

So when (every) year kind of stratifying and looking within one set, one SES category there's nothing poorly defined or ambiguous about the comparison between two providers that is when you start doing the comparison if you're looking at how one provider does in lowest SES patients to how another provider as in highest SES provider patient.

If that point when you're making a judgment about which one is best, you're either making, you know, a value judgment I mean that every provider ought to have the same outcomes no matter what their resource to take or scientific judgment that there's no way possible that the differences could be impacted at all by SES factors possibly but there's a lot of judgment and the only way to avoid those judgments is to kind of do risk case make adjustment approaches that enforce in a kind of apples to apples comparison.

Female: I know we're trying to move on but I just want to say again that there's a lot of evidence that in these cases we may at times be using ways for SES thinking that we're accounting for one thing but in fact it is a marker of the quality that is those patients are exposed to. So you could do the same thing about (inaudible) complication. I mean, choose something that actually is a marker of quality. If you said that this is part of your expected, yes, then you're providing comparing providers with a similar cases, right, you're providing, you're comparing providers with similar for a quality and that's the risks. I'm

not saying that it happens every time. There going to be value judgment but I just want to be clear the risk of (inaudible) not just a value one.

Male: And so if you – in your standardization, there's a problem and there's ways to I did not have that issue by giving issue that differ from convention and direct standardization it's not there in the models that are used by CMS for their readmission measures don't necessarily have that problem but some of our (inaudible) technical but there's …

Female: Great and that's actually right. I think there's a need …

Steven Lipstein: Susannah, there's much evidence. There's actually a lot of evidence to suggest that some of the variation not attributable to differences in quality but difference and attributable for patient population being serve and so if we are going to indiscriminately compare outcomes whether ended up with population you're serving just to say that the literature says that the mostly related to quality and that related patient population just I think different of opinion on that matter.

Susannah Bernheim: Please don't get me wrong. I'm not saying that I what I have said consistently is that's really different for different outcomes. There are plenty of cases on both sides in complex issue. I am only saying that because of the concerns that in some cases the way that the SES plays out has to do with quality rather than in inherent. In some cases we need to be clear in our recommendation because I think we currently are, that you want to take into consideration and not be indiscriminant in just putting it into model. So I totally agree with you Steve. There is lot going on here. This is not simple.

Male: Susannah, I'm wondering if you could share some of the thoughts that you had shared with also about how to actually operationalize the decisions as to whether or not to adjust because I wonder if putting it in, in a little more practical terms might help moving forward here. And I thought you got some, you know, good insights there.

(Crosstalk)

Female: There's an echo on the phone …

Female: I'm sorry.

Female: … if someone could. Thank you.

Female: Is there still an echo? OK.

(Crosstalk)

Female: So now I'm hearing it, too. OK. I will – briefly, I actually thought and I'm very interested in what Sean is saying. I work closely with the statisticians but I am not a statistician and I think it might be useful. I'm happy to raise a couple of issues but I think the idea of cautiously trying to put out some ways that you could operationalize this would be good and we probably can't do it in the next hour on this call.

So I'll say at a high level some the things that we have seen and done and that we suggested that could be I think use as examples to how encasement where there's uncertainty about whether the primary mechanism for SES is through inherent patients factor versus quality factor.

So there are maybe cases where based on as internal literature, conceptual model there isn't a question about that. But in the case where were uncertain, how much the pathway through poor quality is the way that SES is manifesting or meeting to work outcomes versus sort of inherent factors that provider can't influence. You might be able to start this (inaudible). I suspect there's a lot of us on the call who have and could do adjustment as well.

So, you know, I (charter) with the very simplest things. You know, we want to look whether there is at a patient level, a relationship. We want to look at whether that relationship is there after you've accounted for clinical factors so that you –we have found that much of the impact of SES pick by the clinical factors because it has such a – we're looking at Medicare patients and I think it kind of a profound effect on some of the clinical state when they first arrived but it may not be all of the factor that all the effect.

And that you want to look at how providers look at low SES patients do want the measure compared to those without because we have also found in some measures. There almost no difference, and another measure there are the bigger difference.

And then I started to play with some of the ideas that you've seen on literature in other setting just trying the different angle if there are difference among providers as many lowest stratification on those a fewer whether those seemed to be more about the inherent quality of those providers with low SES patients (inaudible) related with a more about inherent patient factor so you can start to this angle so things we have seen that a little bit by looking at the results for the high SES patients within places that have high performance of low SES patient.

We have built some models in this committee just to trying to get some of the stuff that Sean was saying, where we tried to put in both our patient in a hospital factor and I'm learning from my statisticians that you can do this (inaudible) and you can do as well as a reviewing a little careful but then you can start to understand the relative influence of the hospital or the provider compared to the patients.

So, I'm going to stop there. I just think that it maybe it useful to try to at somebody else that give some examples of ways in the cases whether is uncertainty to try and see some of those things out and again, I really support Larry suggestion that in those cases while we're learning about this we try to see some of those things out but that we recommend you provide the score. I think that's a great solution.

Karen Pace: OK. This is Karen. I think good discussion that I think we probably again, need to think about moving on for some other topics and come back to more specific about operationalization. You know, put out have some suggestions that we can circulate and have people refund to that. I think maybe Kevin, do you want to get back to the specifics about disparities recommendations.

Kevin Fiscella: Sure.

Karen Pace: Because I think that may also be the key thing we need to address before we reach the end our time and we can then see if there's anything else that I think maybe the specifics about disparities recommendation will be good.

Kevin Fiscella: Sure. So, hopefully everybody's has a chance to review the appendix that was attached. It was at the back of the memos, so hopefully people got to it. But it's really an attempt to begin to flush out a bit more on the challenge of identifying and really monitoring disparities and potentially the potential unintended or consequences of any actions that are taken as result of recommendations for our panel. And I've put them in to three broad categories.

The first idea was separating out recommendation one and making – identifying disparities in separate recommendations from the first part, because it really seem get lost in the comments. The second one was to begin to flush out the – some of the different functions that stratified data can perform and at what level so, the first one is national monitoring. So, that – I think we would have the opportunity here to begin to actually monitor disparities in these particular measures and then track trends over time.

And then that way, look at, for example the potential that we would have the unintended consequences of actually increasing disparities or hopefully, the opposite making progress and the – these data could actually be shared with AHRQ as part of their national health care disparities report, it would actually enrich that because right now, they are largely based on existing national – federal survey and other types of data, but this would be relevant to the actual quality measure that are being done.

So, that would be a substance of improvement in the status quo. And the second one off course is quality improvement for providers specifically related to closing the gap in disparities by making those data readily available.

And the third one, and this could be an incremental or down the road approach, actually making data available to patients' life themselves so one could – if one had a measure that was adjusted, but provided information for

how a different group perform for patient of similar race or SES, one would be able to access and that's more down the road.

And then the third major suggestion – this would have to the additional recommendation beyond what we already have which would be to have a standing National Quality Forum disparities committee who's task was really to help move this agenda forward both in term of beginning to make recommendations regarding additional measures as well as monitoring the effect by looking at the data on disparities and making recommendations to the National Quality Forum.

So, that's you know, that's the essence of thoughts there. The first and third actually being – what actually made changes in the recommendations that we have because it would mean splitting out the first recommendation, the two recommendations, and the last one would mean adding an additional recommendation for a standing National Quality Forum disparities committee who's task was really focused on the disparities issue including monitoring potential unintended impact.

Lawrence Casalino: This is Larry, I like the second and third a lot. And I also you know, I'm all in favor of stratification and of separated out of recommendation about stratification. By I just want to ask about the status of what were actually trying to say about stratification in this first recommendation. So, it says for purposes of identifying and reducing disparities, performance measure should be stratified.

So, I just would–I'd like to understand that relationship of that too. If we recommend that there be an adjusted measure or an adjusted measure or an unadjusted measure, whatever we come down on that we spend the most of our time today. So, then we're saying how we think the measure ought to be and what relation then of saying performance measure should be stratified to that.

It's almost like telling them that they have a whole another kind of performance measured also. And in a way to me, it's almost like saying – so – I'll just stop there but is this supposed to be in additional thing that should be

done? Is it an additional performance measured? Is it a modification of the performance measure that we're going to recommend? What's the status of this stratification recommendation?

Kevin Fiscella: Well, keep in mind, I guess the first point is, is that it's simply a statement of what the way the current draft reads.

Lawrence Casalino: Right.

Kevin Fiscella: You know, I think the idea would be to have this data available to the public in the same way perhaps the adjusted and unadjusted data would be readily available. But that stratified data would be available for the purposes of identifying disparity because it's been discussed that simply having adjusted and unadjusted don't tell you where the disparities are.

Lawrence Casalino: So with the proposal be then that if you are the measure developer, let's just say we agreed on the adjusted and unadjusted, I'm not presuming that, but just for purposes discussion. So would we be telling that the measure developer then if SES is important, tell us how you'd calculate an adjusted and also tell us how you would measure calculate stratified measures. And so with this be a…

Kevin Fiscella: That's right, that's right.

Lawrence Casalino: OK, so it's not optional. It's at the same rank or importance as saying, "Give us an adjusted and unadjusted measure." is that …

Kevin Fiscella: That's correct.

(Crosstalk)

David Nerenz: … in addition the two parts of the measure?

Lawrence Casalino: I'm sorry I'm still getting an echo, could you repeat that please, David?

David Nerenz: Well, I thought Larry was suggesting that maybe we were imposing too much on the developer if we require on the one hand show the measure both ways.

And then on top of that, we have a recommendation that you have to stratify, do we need …

Lawrence Casalino: And just to clarify, I wasn't suggesting, I was just asking. And I was just suggesting the possibility that some people might think it's might be too much, I don't have an opinion. I mean I like the idea of having the stratified available.

(Crosstalk)

Male: I was to support that (inaudible) stratification that other options should be incorporated.

Female: Who is that speaking?

Male: (Inaudible).

Female: I'm sorry we're getting some feedback …

Male: It's not resolved yet.

Female: OK, I think Nancy …

Female: Nancy Garrett, I think you were trying to get in?

Nancy Garrett: Yes, yes. I was just going to say what I was trying to propose earlier was that instead of requiring that the measure be reported and used both ways to instead (strained) on these various recommendations as Kevin just said. Call it as a separate recommendation, you know, can have some strength to need to run having that data available so that we're actually using it?"

I like the idea of the standing disparities committee and, you know, maybe we could brought in that to also be looking at this whole risk adjustment question because several people have noted, we're just scratching the surface here of this issues, and this format is tough to really get into the depths that we need to. So, I think that those things would really help move us along.

Karen Pace: This is Karen Pace. Just one comment, I think the big issue is as Kevin said whether you provide both a clinically adjusted or totally raw data plus this sociodemographically adjusted measure, the score alone doesn't really illuminate anything about disparity. And so if the real issue is about dealing with disparity then I'm not sure that proposal is really going to satisfy that. So I think that's part of what we want you to think about now.

Female: And just to add up, I just to want to a nod to Marshall Chin has been making this recommendation of a standing committee for a quite sometime. So, thank you, Marshall. We'd put it in the recommendation.

Eugene Nuccio: This is Gene Nuccio. Currently, when a measure is presented to NQF for endorsement, there is a request that we stratify the results based on, on a number different demographic variables, so that the review committee can look at how the measure performs across these various demographics.

Now, is the question whether that information should be made more public or is the question that that sort of Steve was present – perhaps suggesting when he wanted to compare likes to likes, that we create new models based on a particular strata or multiple values of a strata stratification for that purpose. I think that creates a major burden for developers. Developers, I don't think would (balk) very much about have a clinically adjusted outcome and a clinically and sociodemographically adjusted outcome that would be relatively easy for them to do – I mean mechanically.

The issue of presenting it – that information simply to the NQF review committee or making it publicly reportable would – I think is a question that needs to be clarified. So, perhaps we could we could understand – hope you understand whether we're asking for new models for each strata or simply stratifying the results that you get from the two approaches that we've been talking about.

David Nerenz: Dave here. Just a couple of very quick – hopefully, semantic clarifications. Steve suggestion before the break, I would have label this in our, our category of peer group comparisons and in our in-person meeting we spent some time making sure that we had distinguished stratification on the one hand and peer

group comparisons on the other before we're having to do mainly with identifying various sociodemographic groups, either was in a large data set or within even a single plan or provider.

The peer group comparison, the essence being putting like plans and providers together. And we work a bit in January to try to make sure those were two different things. So, I just don't want us to slip back into saying those are two different forms of stratification because they are, they are quite different.

And I guess the other thing just reflecting two or three of the recent comments. Stratification, when we talk about use in identifying disparities can certainly be done in all sorts of different levels by different entities. Individual plans or providers are may desire to stratify their own data on a particular measure based on the groups that are meaningful and numerous and prevalent in their own environment. Purchasers may wish to stratify data across a range of plans and providers according to the set of stratification criteria variables that they find useful.

So I don't know that the issue of stratification is something that falls immediately and heavily on the developers. I had understood our recommendations to be more about the availability of data down stream so to speak for a variety of users, for a variety of purposes getting to do generally with the issue of disparities, not so much with developers had to bring forward.

Eugene Nuccio: OK. Thank you.

Karen Pace: This is Karen. Let me just clarify. We do ask the developers to do what Gene has said but that's more at a aggregate level than a individual provider level which gets sets as function three where you would actually have stratified data available for each provider in a, you know, let's say a drill down from the one score.

Lawrence Casalino: So, this is Larry again. So, this gets back to what I asked earlier. You know, what exactly were asking the, the developers to do so we could ask the developers to give us an adjusted measure and how they going to calculate

that. But what are we asking them to do in regard to either stratification or peer group comparison?

Karen Pace: This is Karen. I think the question on the table is – again, in relationship to this masking disparities or masking differences in quality, the proposal that we spend a lot of time on with presenting two – adjust two scores and the question is does that really get at the issue of masking disparities and the, you know, maybe an alternate proposal is that having the sociodemographic adjusted score in those instances when it really is called for but allowing – but also having the drill down stratified data available. So, it's kind of two ways to look at the issue of making things transparent.

Lawrence Casalino: I understand as a concept but I'm just trying to understand exactly what we're – exactly what we want to – what we propose and to ask developers to do and if we're asking them to present that as an alternative to the risk adjuster or not risk adjustment measures or an addition. I really like to propose how the stratification would be done or how the data – I don't understand exactly what we're asking them.

Karen Pace: Right. So, if these were the recommendation that both of these should be done, the sociodemographic adjustment and stratification, then we – if the developer would be asked to …

Male: Hello.

Female: Dave, we did continue. We're having kind of an emergency here. Sorry.

Male: David? Dave, are you still there?

David Nerenz: Yes. Sorry, I had to set the phone down in a minute and turn away. I don't have a speaker phone. Sorry, I missed the last bit.

Male: I think something happened at NQF and Karen had to stop moderating.

David Nerenz: That's correct. OK. All right. I guess, I'll take up the duties. Sorry, I had to set the phone also for a second but we're still live, I think, with the rest of us.

Male: So, I think Larry's question was what are we asking developers to do here and I think it's as Karen was indicating there's several possibilities. You know, one could be that we just are fully transparent and present adjusted, unadjusted data as well as data that are stratified. And this quite conceivably happened at the provider level so it's like you got an adjusted and unadjusted score. You also got data that unstratified.

David Nerenz: Yes. Let me just – a question I guess to many of us are still here on the call. I have not interpreted the current second part of our recommendation one about stratification to speak primarily a reason (very much) at all to developers. I really here was thinking about downstream users or that was what that message was about.

It strikes me that when developers are putting their materials together, they are working mainly with national data or with other data, whatever is available at the time, they may be able to say – they can do whatever stratification they want to do with that but the stratification that they do is not necessarily the same stratification that any individual plan or provider would do. It's not the same stratification that like a regional purchasing collaborative would do. The (rounded) variables aren't necessary. They're saying that the relevant levels within the variable aren't necessarily the same.

So it seems to me that we're talking about just the general desirability of using stratified data for purposes of attending to disparities but the specifics of how that's done would be quite variable, quite local, depend on the measure, depend on the setting, depend on all sorts of things and it just taught me that that was far beyond what we were asking developers for. But I'm just seeing if others had seen that the same way. The point thing, if we break it up, make it a separate recommendation, I still don't think that we are speaking very directly or heavily to what developers are supposed to do.

Nancy Garrett: David, this is Nancy Garrett. That's the way I interpreted it as well. So maybe, if I (offer) up an example that would be helpful in (inaudible) although we have the measures called the (D5) which is the measure of outpatient diabetes performance and there are five different outcomes that we measure that NQF endorsed actually. And the outcomes have to do with how the

patient is actually how their diabetes is being managed. So it seems like is your blood sugar at the appropriate level? Is your cholesterol in control? Are you a non-smoker?

And so for that measure, right now, there's – we do not do any kind of SES risk adjustments even though patient characteristics and patient behavior does play in a lot to be able to achieve that measure. So for purposes of public reporting and accountability it's, it's unadjusted measure.

I did mentioned that we do have some SES risk adjustment that's kind of buried in the report and it's not being used at all for the accountability and the public reporting, but it's there. It's not very much in the for (finance), it's not used for payments. But my own institution, we are doing a lot of work trying to understand disparities in that measure by raising this in a (SPM) language. And so we are report by – we're doing some reporting to understand by clinic how that varies and by measure. So where do we have disparities that we need to address? How can we use that data to help us improve?

And so that's really not an issue of measure development or endorsement, but that's where – to me, the stratification is essential in order for us to understand disparities and try to address them.

Karen Pace: This is Karen. Sorry for that, we're back. And just to clarify about the developer. I think the way we might think of it is that if, if there's a measure that the developer find should be adjusted to sociodemographic factors based on all of the caveats we've talked about. And obviously, they would be in the best position to identify what the stratification categories would be. But they wouldn't be the one responsible for implementing stratification. That would again come with the implementation.

So, I think that the key question is whether there is, you know, we want. I mean it really is about dealing with these disparities concern whether the approaches to say we suggest both, you know, clinically adjusted and sociodemographically adjusted scores or whether we really talk about in the case that there's sociodemographic adjustment that the stratified data would be available along with that sociodemographic adjusted score.

Steven Lipstein: So, Karen this is Steve again. When you – can you add some clarity to the stratification recommendation in that – one level of stratification is you begin to compare communities like – with communities of like characteristics and like, perhaps, sociodemographics. The other part of this is that the local circumstances are highly germane to patient outcomes in terms of how local communities are (resourced).

And we talked about this at our face-to-face meeting where a community like (I believe) community like Chapel Hill would be resourced very differently than resourced like the in the community like the Boot Hill, Missouri and yet you wouldn't be able to tell that by just looking into the (med card) data set. So, how could it, how can we take a recommendation that suggest the stratifications. It has to be – you need to strengthen the data available to really make the strata meaningful in terms of comparisons.

Karen Pace: Yes. And I think and Dave was saying that the – what you're referring to is like comparisons is what we referred to as peer group comparisons. And that that's tend to be more in the implementation side than how the measure would be specified. So, I think in this case as the original stratification recommendation and what we're continuing to talk about here had to do with stratification by those factors which could be at the provider level, it could be at the community level, it could be at a national level.

But what you're mentioning is what MedPAC recommended maybe on a different variable. But, you know, we've been referring to that as peer group comparison. And I think where that is in our current recommendation is that what's – for those who are implementing to really consider – even if we do sociodemographic adjustments, you know, they still need to consider other avenues when they're putting these into programs of whether additional approaches are needed on top of that.

Lawrence Casalino: So Karen, this is Larry. We're hearing two different things and I think it – you probably didn't hear well because I think you had to go. I'm glad you guys are safe.

I mean at least some of the committee seems to have any impression that the discussion of stratification or appear for comparison is more along the lines of you might want to do this if you're provider, or health plan, or whoever. And it has nothing to do with the developer.

And then other people seem to think that now the developer or to say something about stratification. So, you know, if it's the ladder, I mean then here's the question if they "say something about stratification" or are they proposing or peer comparison or they're proposing a measure and this is a measure that we then approved or not by NQF just as a proposal to have an adjusted or unadjusted score would be approved or not approved.

These are two quite different things. And I'm not sure some people say to seem to be turn that one, some people say seems to be turn with the other.

Karen Pace: Right. And I guess what I'm saying is I think we'll be come back where this panel ultimately lands on what's the best approach to – in addition to sociodemographic adjusted score, whether additionally the unadjusted score either totally unadjusted or just clinically adjusted score is also available. Or whether the recommendation is it should be a sociodemographic adjusted score plus the stratified results.

So, I kind of seeing them as different options and depending on the recommendation and what ultimately gets approved by NQF would then lead to what the developer had to do, if the results was – it should be stratified in addition to the sociodemographic adjustment, then we would ask the developer to identify what strata should be. And they would do some analysis that way.

So, I don't think we should start what the developer would be required that more depend on ultimately where we end up with this recommendations.

David Nerenz: And Dave here. Two things (to respond) to Karen. I don't want to belabor the point. I've been comfortable all along with the wording of what we had in the report and, you know, happy to see it broken out separately.

I like the wording because it struck me is that it allowed – it didn't prescribe – for example, what the strata should be. And it didn't even presume that such standardization could be done. I guess that's the point I make now.

It seems to me that defining of the strata how they were caught, which variables, which groups are very a local decision, they're application dependent, they're region dependent, they're measure dependent, they're beyond measure dependent, they're circumstance dependent.

I'm thinking for example project we did in number of years ago about disparities in (mammography) rates in the health plan. We ended cutting – eventually we had 30 different cells which is a product when you think of four different stratifying variables, but there's not necessarily reason to think that anybody else we cut the same way and choose to do it that way.

So it seems like a general recommendations that the value of stratified data, the positive value then approach with disparities, that's great. I think that's what we said. But I do wonder a bit about what we can ever require or NQF require developers about saying what the strata should be, I don't know how can they do that.

Lawrence Casalino: This is Larry. I agree with that. I think …

(Crosstalk)

Male: Go ahead.

Lawrence Casalino: … which (inaudible) might be very useful but as a alternate requirement for what the developer has to do, I'm not sure that it could be done well, for the reason that Dave – another people just given.

Karen Pace: OK.

(Off-mike)

Karen Pace: David Hopkins is that you?

David Hopkins: Yes. Just agreeing with what – I think it was Dave Nerenz was saying.

Karen Pace: OK. All right, did we ever get with Marshall Chin? Is he still on and want him make any comments in this area? OK. We need to open the line for public comment. So I know that this is been an interesting conversation, a lot of good ideas and we will need to huddle after this to see how to lapse things up with you or proceed further with discussions.

Let's first see if there's any public comment. And operator would you facilitate that please.

Operator: Thank you. At this time, if you have comment, please press star then the number 1 on your telephone keypad.

You have a comment from John Shaw.

Karen Pace: OK.

John Shaw: Hi. This is John Shaw from Next Wave up in Albany. I very much enjoyed the discussion today and I think we're very much on the right track particularly some of the specific recommendations to report the two measures the adjusted for SES and the clinical (only) adjustment.

And part of the reason for that is I'm a measure user. And what I care about measures how can it help achieve better outcomes. And so, if there's variability there, if we see poor outcomes, we need to do something. And what having both of those measures available will help us do if to see where and what is the primary control point. So is it inside and better clinical interventions and better clinical quality, or is it outside something that's affected by the patient and their informal caregivers in the community and all of the social interventions that are there.

We're moving towards getting away from free-for-service and getting more patient-oriented health and wellness. And so a lot of the controversy back and forth is how we pay for things and we're starting to see some light at the end of the tunnel that would allow payment for both inside and outside social intervention as well, through ACOs, (cellphones). In New York several of us

are involved in implementing the Medicaid waiver to adjust this. And I'm thinking how might I use the tools to do that.

One of the things that might help with visualizing peer grouping stratification and addressing the disparity is what I intend to do if I can get those two measures is immediately take the ratio. What's the ratio of the SES risk-adjustment measure versus the unadjusted, and that gives me a measure of degree of SES challenge either plus or minus, either I have a provider that's serving primarily and affluent, well-educated population or one that's challenge. And that might give the providers a way to simply find similarly situated peers to compare to by just looking at the values of the ratios.

The other thing I wanted to note is depending on what the measure is the drivers for SES maybe very different depending on the measure. So we're focused in Albany on asthma, COPD, diabetes, and mental health issues.

And if I looked at what we're doing in terms of designing internal and social interventions, they're very different in those populations, for example with asthma and COPD, the home and school environments play a big factor where they might not elsewhere.

So, I think keep going in that direction, let's get some information out there and additionally give us the opportunity to drill down into the data for whichever the measures to take it one step further at the local or national level. Thank you.

Karen Pace: OK. Thank you. Operator anyone else?

Operator: Yes, you do have a comment from Nancy Foster.

Karen Pace: OK.

Nancy Foster: Good afternoon. Thanks Karen. Good afternoon to everybody on the panel and appreciate this opportunity to listen into your discussions and to comment.

As you saw from our comment letter, we were absolutely very much in favor of the recommendations and you originally stated them, like I say that

wouldn't be in favor of any changes that you might make. But we do believe you have correctly identified an issue that very much needs to be addressed through the NQF process, and that is essentially the NQF current policy as we understand it directs measure developers to essentially not think about a whole series of factors that may influence outcomes or perhaps other measures but principally outcomes.

And n that sense it really is very limiting. If my colleague's article and other articles that have recently been written about the readmission and other measures are correct, there are in fact a number of factors that affect the likelihood of readmission. There's just one of those outcomes including the quality of the nursing home care that the patient has sent to the prevalence of primary care physicians in the community, whether or not the patient lived alone.

All of these are factors that I think really warrant attention and understanding and by limiting our ability to call them out in the measurement world. We are really blinding ourselves to the impact of those factors and limiting our ability to address these really underlying causes and drivers of disparities and outcomes.

So, I've very, very much appreciate the comments and the challenges that you've been presented with and trying to outline all of these for the National Quality Forum and I hope that you will continue to strongly favor illuminating our understanding of what are the drivers of disparities and outcomes.

If it happens to be the quality of care being provided, then we need to know that and we need to fully understand that. If it's other factors, I think as a country, we need to understand that and begin to effectively address those others factors as well.

So, thank you for you work.

Karen Pace: Thank you, Nancy. Operator.

Operator: Your next comment comes from Tom James.

Nancy Pace: OK, go ahead Tom.

Tom James: All right. Thank you. Good afternoon. Tom James with AmeriHealth Caritas. Just a couple of quick comments. We've very much appreciate the direction that you're moving. I spoke with the FQHCs in your area the Federally Qualified Health Centers. They're very much looking forward to having these kinds of measures because they feel that they are really trying to provide the high level of quality and it appears in the standard types of measurement.

And by the same token, two weeks ago when I was on the cardiovascular workgroup, we identified several measures that we wish we could couple with work that you are doing.

And then as a final point, I think this is going to your work that you're considering so carefully, is important because so many Medicaid plans as well as Medicare plans, these are risk-adjustment methodology, it is applied using only medical factor. I suspect that with your work if this will move into the socioeconomics factors that will do further risk adjustments for Medicaid and Medicare going to the future.

So, please keep up the good work. Thank you.

Karen Pace: Thank you.

Operator: And there are no further comments.

Karen Pace: OK. So Dave and Kevin, I wonder if you have any suggestions for how we proceed one might be just asked the panel is there any key issues that we didn't get to that we need to pay special attention to. If you have other thoughts, welcome.

David Nerenz: Dave here, that would certainly be one given that we're – just a few minutes away from hard stop. I think beyond that what I think we could commit to doing is to putting out to the group basically a set of proposals based on the discussion today about tweaks, changes addition and whether immediately or certainly they're after go to the same kind of process we went through earlier

about support, can live with, don't support because I think we've got, you know, by my counting maybe five, six, seven things that we have discussed today. We can't really do show of hands in a phone format but I think we can move quickly after the call.

Kevin Fiscella: Yes, I think that's a great plan. If people feel that issues have – that it didn't have a chance to raise, please, you know, e-mail either us or the entire group if you feel warranted.

Karen Pace: So we'd like to hear from anyone that things anything in the memo or anything in the comments that didn't get in the memo that you want us to pay special attention to and get back to you on, you know, as Dave said – with Dave and Kevin, we'll get something address put together for you to respond to.

Male: So Karen if I hear correctly the process from here will be summarize some of the key suggestions made today, get everybody's input, some sort of scoring mechanism. And then take the results of that and we actually have some consensus do some kind of modification of the report. And then will be there an opportunity to look the final report.

Karen Pace: Yes, that would be the plan. And I think in terms of just logistics, obviously the most important thing is to get the recommendations correct. We know there are suggestions about improving the report and the language in the report and we will continue to work through those.

But I think our first priority is to make sure we have it solid on what the recommendation from this group that will move forward to the SCAC. But we will definitely – you know, we've already have some thoughts and obviously we have suggestions on improving the report and we can continue to do that.

But the first priority is the actual recommendation.

Kevin Fiscella: Yes. And that's actually what I meant when I said that. So …

Karen Pace: OK.

Kevin Fiscella: Is it that the next thing we're going to see is a revise set of recommendations following this discussion? Or there are a few key points if I heard Dave. He's thinking that you would want to pull us on …

Karen Pace: Yes.

Kevin Fiscella: … and then reduce recommendation. So it is still two steps.

Karen Pace: Right, it's still two steps. Thank you.

David Nerenz: Make sense.

Female: Thanks for your four hours of engagement everybody.

Karen Pace: Yes, definitely e-mail us, phone call us as your thinking of things because we'd love to hear that. And thank you all so much for devoting this time today on the Friday afternoon – morning and afternoon. So we appreciate that and we will get back to you as quickly as possible early next week with some things to react to. OK.

David Nerenz: One last bit of thought again, thank you so much for the way in which all of you have approach this task and the interactions, the commitment, the respect and collegiality. I think this is going to wonderful process of exchange and learning from each other, it's been really good.

With that in mind, let me say, I've just valued (the) pleasure the level of consensus we've able to achieve and the common ground we've able to find. So as we move from here to the finish line, our goal is certainly not to do anything that we move away from the large amount of common ground we have already. I think – or at least speaking for myself, personally, my focus is going to be on those things we discussed today that would create an even higher level of consensus than we currently have rather than things that are going to splinter in some way.

Female: Great.

Karen Pace: Kevin, any final words or let you …

END