NATIONAL QUALITY FORUM

+ + + + +

PATIENT OUTCOMES ALL-CAUSE
READMISSIONS STEERING COMMITTEE

+ + + + +

TUESDAY
DECEMBER 6, 2011

+ + + + +

The Steering Committee met at the National Quality Forum, 9th Floor Conference Room, 1030 15th Street, N.W., Washington, D.C., at 8:00 a.m., Sherrie Kaplan, Chair, presiding.

PRESENT:
SHERRIE KAPLAN, Chair
TANYA ALTERAS
BRENT ASPLIN
RICHARD BANKOWITZ
JIM BELLOWS
JO ANN BROOKS

PAULA FOLTZ
FRANK GHINASSI
LAURENT GLANCE
JEFFREY GREENWALD
BRUCE HALL
LESLIE KELLY HALL
ASHISH JHA

MICHAEL LANGBERG
ELIOT LAZAR*
PATRICIA McDERMOTT*
DAVID POLAKOFF
BRUCE POMERANZ
MARK SCHUSTER
CHRISTINE TRAVIS

NQF STAFF PRESENT:

TAROON AMIN

HEIDI BOSSLEY

HELEN BURSTIN

JANET CORRIGAN

ALEXIS FORMAN MORGAN

KAREN JOHNSON

ADEELA KHAN

LAURA MILLER

KAREN PACE

ALSO PRESENT:

DAWN ALAYON, NCQA

ELIZABETH DRYE, Yale University

NANCY FOSTER, American Hospital Association

JEREMY GOTTLICH, NCQA

JEPH HERRIN, Yale University

LEORA HORWITZ, Yale University

RABIA KHAN, CMS

HARLAN KRUMHOLZ, Yale University*

KAREN NAKANO, CMS

MARA RUBIN, UnitedHealthcare

ROBERT SAUNDERS, NCQA

GRAEME SCANDRETT, UnitedHealthcare

RON STETTLER, UnitedHealthcare

*Participating via telephone

C-O-N-T-E-N-T-S

P-R-O-C-E-E-D-I-N-G-S

8:43 a.m.

MR. AMIN: Okay, so I will begin by outlining the discussion format for the next two hours. The discussion will begin with the CMS Measure 1789.

The Chair will request that the measure developer will provide a 10-minute response to each of the requests for further information discussed by the Committee yesterday.

For CMS, these requests include, as a preliminary start, although if there are additional questions feel free to ask, details on the distributional properties for small hospitals using HLM, the use of the following covariates in the risk model, including SES and hospital volume, and additional concerns raised this morning, calibration curves, risk decile plots and the usability and use of this measure for quality improvement.

The Committee will then have 30

minutes to discuss and ask further clarification questions of the measure developer and then decide whether to move on to a final vote on scientific acceptability and usability and an overall vote of the measure.

We will use the same format for the NCQA measure.

CHAIR KAPLAN: I would just like the developers -- welcome back and thank you very much for being so responsive, and we have 10 minutes you are going to get for an uninterrupted presentation and then we are going to have an additional 30-minute discussion with you about the range of topics that you know, that Taroon just outlined, in addition to any other FOCUSED concerns, all caps, that the Committee has.

MS. HORWITZ: Thank you. We really appreciate the opportunity to come back. So I am going to begin with calibration, and let me just, before I show

you the slides, let me just orient the group to what this is.

So the calibration question, and correct me if I am providing something that you don't want, is a question of how well the model is able to predict risk for aggregate groups of patients.

So if you are looking at low-risk patients, how well is a model able to predict their low risk, and if you look at high risk, how well is the model able to predict that.

And in the ideal world, we would plot the observed readmission risk for an aggregate group of patients versus the expected, and then -- or the predicted, and we would see a 45-degree line. It should be perfectly aligned.

And last night I couldn't quite figure out how to convert our tables into exactly that graph, and so I'm showing you a slightly different graph, the same idea, which is both the observed and the predicted plotted

simultaneously.

So let's show the first graph and I'll show you how this is shown. So on the bottom here we have risk deciles. So on the far left we have the lowest-risk patients and on the far right we have the highest-risk patients in deciles, and on the left we have the readmission risk.

And what we want to see here is not a 45-degree line, because we are not plotting it in that way. What you want to see here is overlap of these two points, of the observed and the predicted, as much as possible, and ideally they should be perfectly overlapped.

So this is our medicine cohort, and you will see they are very, very close, all the way through, with a slight overprediction of risk in the higher end, lower end, but really pretty close.

And if you look to the next slide, you will see our surgery/gynecology. Again,

a slightly bigger spread up at the high end, but generally speaking pretty good, we think, risk prediction.

Similarly in the next graph you'll see our cardiorespiratory, where it's identical, really total overlap. I swear to you that there are purple things -- or blue things underneath there.

And in our next, we have cardiovascular. Again, slight overprediction at the high end, so we are giving people a little more credit for having higher risk patients than they really do, but pretty close.

And in neurology, again very close. So we would make the contention that each model performs quite well on our calibration.

So the next topic that you asked us to talk about was the issue of small volume hospitals and hierarchical regression and what this all means.

And so we said that we would come back today with some data about how the small volume hospitals perform and so that we could put some data around this discussion.

So if you turn to the next slide, this is a distribution of volume, just so you get a sense as to what we are talking about in our data set. Remember, this is all conditions. This is not just single condition.

And you can see that our median volume here is 750 patients. These are big volume observations or hospitals.

We sort of arbitrarily decided we would define small volume for the purposes of today's discussion at the 10th percentile, so we will just take the 10th, you know, the smallest 10 percent of hospitals in our data set, and even those have 88 or fewer patients. So even our 10th percentile, at 88, is quite a lot.

And you will recall that we don't

publicly report anything below 25 currently. So the number that's actually between 25 and 88 is only about 200 hospitals. It's a very small number that have even this very small number of patients. And this is again about 5,000 hospitals altogether.

Now, we weren't sure if you were thinking about small hospitals in terms of the actual hospital itself, so we also defined it for today's purpose by bed size.

So if you look at bed size of the hospitals. Now, this is using American Hospital Association data, so we don't have data on every hospital, but we have data on 4,700 of them.

And you'll see again the median hospital in this data set has 100 beds and the 10th percentile and smaller hospitals have 25 or fewer beds, so these are very small hospitals, hospitals with 25 beds only.

And again there's, you know, about 400 or so of these, this is 10 percent of the

sample.

Okay, so let me show you results both for hospitals, small hospitals as defined by volume, and small hospitals as defined by bed size.

So in the volume, once we kick out the hospitals that don't even have 25, we are left with 292 hospitals that have 88 or fewer patients a year. So, again, less than, you know, a couple a week.

And you'll see -- I don't have an arrow -- but you'll see first of all the medians are the same as you would expect, and that the range, so let's take say the 10th to the 90th percentile, the sort of distribution of it is quite wide, slightly narrower than the overall data set, but not that much narrower, and still a pretty substantial range in performance.

And you'll see this even more dramatically with 25-bed or fewer hospitals -- all right, so we have got -- there it is. So

we've got, you know, if you look at the extremes around 10 percent to 90 percent, they are slightly different but relatively similar, and in bed size, when we are taking less than 25, here we have more hospitals -- 400 of them, still about 10 percent of our sample, the smallest 10 percent -- and here you'll see that the 10th percentile and 90th percentile are again very similar.

So we still have wide distributions of performance even in -- or similar distributions of performance even for the small hospitals versus the large hospitals.

So we would argue that what we are not seeing is a very tight pull to the mean that people are concerned about. So we are not seeing the sort of 10th to 90th percentile all tucked right around that median in the middle which is, I think, the main concern that the Committee was expressing.

Despite the fact that we don't

empirically see very much to be concerned about, I want to just take a minute to explain a little bit what our modeling approach does, because we didn't really have a chance to explain the issue and I think that there's some confusion about what it is that we are doing with this model.

I want to first say that this is a widely-used, pretty standard approach. It's used in the NQF-endorsed public report measures that we have now for the three conditions.

But it's also used in many, many other NQF-endorsed measures as well as national registries in Canada and the UK and elsewhere. So this, in many respects, is the standard approach to modeling.

There's two things that this approach does. The first thing is that it takes into account the clustering of patients within hospitals.

And what we mean by that is that

most statistical models assume independence of samples. They assume that the -- any observation's likelihood of an outcome is totally independent of some other observation's likelihood of an outcome.

And we know that that's just not true when -- for this data set. Our data set violates that assumption, because we have patients within hospitals, and we believe there's a hospital effect, just like we use the same approach for looking at education, where we are trying to assess student performance, we know that student performance is affected by their teachers' quality, and so the same thing, we use the same kind of modeling approach to look at teacher quality.

So because we know that the samples are not independent, we have to account for the standard error that we get, and this model accounts for that.

Now, there are other ways to do that as well. This is not the only way to do

that. But that's one advantage of this approach.

The second thing that this model does is it takes into account the increased variability in small hospitals as well as our knowledge about them up front.

And I want to explain that by way of an example. So if I take a quarter out of my pocket and I flip it five times and I get four heads, then I have a heads proportion of 80 percent, okay?

Now, I'm going to flip it again. And before I flip it again, I want to ask you, what do you think the likelihood of heads is going to be?

Now, in a traditional approach, even one that accounts for clustering and even one that accounts for the hospital effect, we would say that, well, the data that we have suggests that it's 80 percent, it's a small volume, so you know we have a wide confidence interval, so we'll say oh, it could be

anything from 50 to 100, but with a central tendency of 80. But nobody here would take that back, right? No one here is going to give me four to one odds that I'm going to get heads when I flip this coin again, because we have some knowledge about coins.

Right? We know that 99.9999999 percent of coins are fair, so we expect that when we flip this coin again, actually what we are going to get is 50-50 heads.

So which estimate is closer to truth in the universe, the 80 percent that we get just from the data, from the sort of standard model, or the 50 percent that we get from pulling it towards what we know a priori about coins?

And that's what our model does. So our model accounts for the fact that when we have a lot of variability, and when we have small samples, and when we know something about that sample, that it's in fact more accurate to use that prior assumption of what

we know about that process to understand what the estimate is likely to be, and that's what our model accomplishes.

So if I flip this coin 500 times, and I got 400 heads, and you asked me what I am going to get the next time I flip it, I'm not going to say 50-50, right?

Because even though I know something about coins, now I have a very large sample and I am much less inclined to think that this coin is fair.

So the larger the sample you have, the more you can trust the data that you have, but in a small sample, it's more accurate and more fair to consider that prior probability.

So that brings me to the next question that the panel asked us to address, which is, well, are small hospitals hospitals? So should we treat our a priori assumption as that they have a hospital effect? Or are small hospitals small hospitals, and should we consider instead that their a priori

assumption is that of the small hospital?

Therefore should we include some kind of volume indicator in our model? And that's an interesting question and a sort of policy question.

I will say up front that we don't have data to suggest that small hospitals in aggregate behave differently from large hospitals for the outcome of readmission.

We haven't looked at it carefully in our data so I can't speak to our data. We are not aware really of solid evidence about this.

Whereas there's a very good evidence base around volume outcome for mortality or for complications, for readmission we don't have that data.

So the first thing we would say is we are actually not sure that this is a problem, and it's not obvious that it should be a problem necessarily. Small hospitals might do a pretty good job of taking good care

of their patients and doing medication reconciliation because they have fewer patients to handle.

So it's not intrinsically obvious to us. And second, even if it were the case, it's not clear to us that all small hospitals are the same and should therefore be polled that way.

So we have leaned up until now not to including volume as a separate indicator, not to considering a separate up front assumption for small hospitals, and we'd rather just level the playing field and say look, we are going to treat small hospitals as hospitals and we are going to assume that their baseline is the same baseline as all other hospitals until proven otherwise by data.

MR. AMIN: Not to interrupt, but I just want to say you're at 10 minutes, but I know you have a few other slides to go, so we will just give you five more minutes to just

finish up and there will be time for a back and forth. But we just wanted to make sure.

MS. HORWITZ: Great. So our next slide is about SES, which is the next topic that you asked us to address, and I want to just show you again the data in our data set about how hospitals perform in terms of their proportion of dual eligible patients.

So we divided our data set sort of a priori into four quartiles -- less than 10 percent of their patient -- of a hospital's patient population having Medicare and Medicaid, all the way up to more than 30 percent of their patients being dual eligible.

And what we wanted to show you were two things -- I am going to use this crazy thing again here -- okay, so first of all the median result here.

So, the median result for the hospitals with the fewest Medicaid patients is 16.5 and for the hospitals with the largest number of Medicaid patients it's 16.9.

Yes, there is a relationship. It's very small.

The second thing is that the hospitals with the highest proportion of Medicaid patients, a quarter of them perform better than the average hospital with very few Medicaid patients.

So the overlap is substantial. There is not necessarily a defining sort of destiny around the proportion of SES patients that you have in your hospital, and that's one major reason that we were not that inclined to try to include SES in our models, because, in fact, hospitals can perform very well with a high proportion of the Medicaid patients, as I think, the State Medicaid Directory in Massachusetts demonstrated yesterday.

The second thing around -- maybe I'll just leave that as it is for SES and I will just comment lastly, because you wanted us to comment on usability.

So there were two, I think, main

usability concerns that the panel raised yesterday. One was a question of usability for patients, payors, consumers, around the question of, well, you know, will they be able to use data around small hospitals, will we be able to find a difference, or is it just going to be all these hospitals we can't tell?

And I hope that I've demonstrated to you that that's just not a concern for the vast majority of hospitals in our data set, that there are some hospitals with very, very few patients for whom we do pull the results a little bit towards the mean, but I hope that has convinced you that that's actually an appropriate way to estimate their risk rather than a bias.

There was another concern raised about how patients and payors could interpret Hospital Compare data, and I will only comment that it's hard to present data in confidence intervals and it probably could be done better, or potentially it could be done better

on Hospital Compare, and that would be great.

But that's not a function of the measure per se, that's just a function of the way you choose to report the results, which could be reported in many different ways and probably could be done better.

And lastly, I wanted to comment that there were questions raised about usability for hospitals, and here I want to comment a little bit about our intention in making the measure also from a personal perspective because I chair the Readmissions Committee at Yale, and we are a high outlier for many of these conditions so we are on the one hand making these measures and then penalizing ourselves for them, and so it's my job to fix that as well.

And so let me just comment that we built this measure for two purposes. One is for public reporting and public reporting is a heavy responsibility. You have to be really convinced that you are doing it well and doing

it right, so that you are adequately able to compare different types of hospitals, and I think you have all seen and agreed with our risk adjustment and our strategy for defining our outcomes, and our inclusion criteria have really been carefully built to make sure that we are appropriately able to compare hospitals.

But the second reason we built this measure was for quality improvement and what we intend this measure to do is to allow hospitals to benchmark themselves against other hospitals to identify areas in which quality improvement is necessary, and then to catalyze activity.

And in fact we not only built a hospital-wide outcome but because we have five cohorts we can actually provide detailed data on each of these service lines without a hospital, which increases the usability for a hospital in terms of figuring out which areas it's bad -- it's sort of not doing well in.

But what we did not build this measure for was for use in rapid cycle interventions within a hospital. Now I, as the readmission chair, am very preoccupied with my real-time readmission rate, and we track that monthly. We track it across the hospital. We track it by unit. We track it by condition. We are very vigorous about tracking it. We do tons of interventions. We want to know how well they are working.

And we don't need a risk-adjusted measure for that because my case mix doesn't change that much month to month. My distribution of conditions doesn't change much.

In fact, my raw readmission rate at Yale-New Haven Hospital, I regret to say, has been rock solid stable for three years. It hasn't budged by even as much as 0.2 percent. It is exactly the same.

So when you are using data internally to track your own results and your

results over time, you don't need a risk-adjusted measure and that's not what we built this for.

This is really so that we can identify, do we have a problem at all, and what kind of a problem do we have?

MR. AMIN: All right, thank you. Thank you very much.

CHAIR KAPLAN: Okay, now I would invite the Committee to have a 30-minute discussion along the lines that we discussed in executive session and we are open for comments and questions. Ashish?

MEMBER JHA: So, first of all, thank you guys for doing all of that extra work on incredibly short notice. I can't imagine what your evening and morning were like.

But I know that the Committee is really grateful that you guys did all of this. So thank you.

If you can go back to the slide on

volume for a second, or size.

So I want to bring up the example that you used with the coins, and you flipped the coin five times and you got four heads. Now, if we all got a coin and we all flipped it five times and we all got four heads, one would not, I think, at that point -- let's say there were 200 of us in the room -- you wouldn't begin by saying that the central tendency here is 50 percent.

You would say there's something systematic going on. These coins are funny. And you'd need empirical evidence to show that.

So no one, I think, has ever suggested that you take a single, small hospital with five observations and do anything with that. That would be silly.

There is actually very good evidence. We had a paper in Annals of Internal Medicine this past year. We have basically done -- and it's out in the

literature so this is not unpublished work -- so the paper in Annals of Internal Medicine showed a very nice volume readmission rate relationship.

We used -- and we have done this, and other people have done this using identical risk adjustment models to what CMS uses. We just don't shrink.

And to go back to the coin example, if you individually shrink each coin five flips, everybody comes out at 50 percent, then when you aggregate it, it looks like, hey, these coins are at 50 percent, when we know that if you had not shrunk and looked at them in aggregate, there was something going on.

What's really interesting to me, when I look at the numbers you've pointed out, so you didn't run the numbers without shrinking. I understand.

But when I look at those, first of all, what's interesting is that small

hospitals have slightly lower readmission rates than large hospitals, that the upper and lower quartile performance for the small hospitals is better than the large hospitals, and the 90th percentile and the 10th percentile performance is better for the small hospitals than the large hospitals, and that's really interesting.

And then when you think about who the outliers are going to be, what happens when you shrink it in the way that you choose to, is that if you worked at a large hospital, you're far more likely to be labeled as an outlier because the data make it so that you'll -- so the small hospitals can pretty much never make it as an outlier, or your performance has to be dramatically worse in order for that to happen.

And so it is a fairness issue but it's also, in my mind, how you use the data, and I guess the question is: what's the philosophical problem with using past

performance?

I understand that volume raises a series of questions, but what's the problem with saying, in the prediction model, if you have been -- let's say your readmission rate has been 35 percent for five years in a row, we are not going to begin, in year six, with the assumption that you are at the national average again.

We have five years of data suggesting that you are where you are. Is there a mathematical problem with that? Is there a philosophical problem with that?

Wouldn't that be more friendly to consumers because it would actually use the information that's available?

MS. HORWITZ: Shall I respond to that? These are important questions, and to a large extent these are policy questions, in terms of how you want to handle hospitals and how you want to think about leveling the playing field.

I want to just first correct one thing, which is these are HDLM results so it's very hard to say what small hospitals are actually performing. You can't use these data for that question, now, because we have already pulled -- we have already assumed an a priori.

So I don't really know what the results would look like had we not done that.

So these are hierarchical regression results, so we have already done the procedure that I talked about in the sense of assuming the average hospital and it's pulled it in that way, so we don't -- but you were commenting on some differences and it's hard to know if they would be worse or if they would be better without that.

The second thing I want to say is that if we -- let's suppose we had picked a different mean to pull small hospitals to.

So we would say, okay, you are a small hospital, your average is going to be

slightly higher, say, than the hospitals at large.

So what does that do? That, first of all -- so now you've got kind of a bumpy playing field instead of a flat playing field -- and we are saying okay, if you are a small hospital who is performing very well, we are now going to pull you higher because your average small hospital performs worse and we make it probably virtually impossible for a small hospital to look like they are better than average.

Conversely, we make it much easier for a small hospital to look worse than average because we are pulling them in that direction, and it remains an open question to us whether that's the appropriate measure to take, given that we don't really know that all small hospitals are homogenous in that way and that it is sort of appropriate to change the playing field for them.

Now this is a hospital-wide

measure we are talking about, not a condition-specific measure, so this is not a sort of condition -- whether there's a volume-outcome relationship for specific conditions is different from whether that's true at the hospital level at large.

CHAIR KAPLAN: Okay, what I would ask, though, Ashish, is that we keep the questions concise and the responses concise. Thank you.

MEMBER JHA: I'll reiterate my question, which is there are lots of ways of doing it and we don't need to get into a discussion about that.

How about the hospital's past performance as -- it's pretty fair, and you could argue that it's probably the best predictor in aggregate of a hospital's performance, is its past performance.

MS. HORWITZ: Yes, so, you know, I think that that's a reasonable approach. We have not looked at that so I can't tell you

how it would change the outcomes or how it would change the measure.

CHAIR KAPLAN: Thanks very much. Other questions? Bruce?

MEMBER HALL: I just want to reiterate Ashish's thanks for all the hard work and I am going to ask a slightly different technical question.

Your measure represents a volume-weighted geometric mean of these cohorts. Do you have any concerns about individual hospitals still being dominated by a particular cohort and whether that would lead to any sort of out of sample comparisons, where one hospital's evaluation is still dominated by one particular service line, and then you are creating an impression of perhaps more standardization there than there really is?

MS. HORWITZ: Well, we debated this a little bit, and our feeling is that we want to represent what hospitals are actually

seeing, so that because we volume-weight, we are allowing the patients that hospitals really see to dominate their measure.

And so if the hospital is dominated by one cohort, that's because those are the patients it's predominantly seeing, and we think that's actually appropriate.

MEMBER HALL: So in effect what you are saying is, if any other hospital in the country had to do what we have to do, this is what the expectation would be?

MS. HORWITZ: Yes, so, for each cohort, we are comparing your performance on those -- the patients in that cohort to all other hospitals' performance for patients in that cohort, and then we weight your performance on that cohort by the number of patients you have seen in that cohort, to give you your average result.

CHAIR KAPLAN: Thanks. Laurent?

MEMBER GLANCE: Thanks again for doing all that work. A quick question, very

quick question. For your calibration plots, were those an independent data sample?

MS. HORWITZ: Yes, so obviously we can't do it on the derivations sample because then we have perfect, that by definition is perfect.

So what we used -- we have a split sample, 2007/2008 combined. We split them randomly in half. These data that I am showing you are from the validation of the '07/'08 validation sample.

We have similar data for 2009, which I am not showing you, but I will tell you that they are very similar.

CHAIR KAPLAN: Richard.

MEMBER BANKOWITZ: Again, thank you for the presentation. In the data on the socioeconomic impact, I have heard a couple of explanations that -- one is that, well, there are other comorbidities involved with higher Medicaid populations, and so what you are showing us is not -- it's not possible to

discern the marginal effects of this.

To do that you'd have to included it in the model and see. Do you know if there are marginal effects to the Medicaid population, and if so how would you handle those marginal effects?

MS. HORWITZ: So you're asking if we added some kind of measure of SES to our model, would our model perform better?

We didn't do those analyses and for two reasons. One is, it's actually very hard to come up with a reliable and acceptable proxy for SES using administrative data. We are not really aware of any measure that we think has appropriate validity.

The second reason is that, let's suppose we put that in the model and it did improve the performance of the model, we are not sure how to interpret that, because we don't understand still, we can't disentangle what proportion of real outcome difference that's attributable to SES has to do with

intrinsic characteristics of the patients that are just totally unavoidable, and what proportion has to do with quality of the hospital, bias, or the ability of the hospital to handle these patients, or the type of health literacy materials you give, or the type of social support or the community relationships you build, and so on and so forth.

So even if we found a relationship or an improvement in the model, it's not clear to us how we would interpret that.

CHAIR KAPLAN: Thank you. Other questions? Ashish.

MEMBER JHA: I have a quick interpretation. Can you go forward a few slides? It was something you guys popped up yesterday showing the distribution. I just want to make sure I understood some of the numbers.

There was a slide you showed yesterday that basically showed the

distribution of the performance, and it's in the technical --

MS. HORWITZ: It's in the technical report, if you have that handy.

MEMBER JHA: Right. What I remember is the difference between the 10th and the 90th percentile -- and I'm going to make up the numbers because I don't remember -- it was like 15 to 18 percent, suggesting that 80 percent of hospitals are between 15 percent readmissions and 18 percent readmission or a 3 percent gap between the best performers and the worst performers if you think of the 90th and 10th percentile.

First I want to make sure I got that right. I understood --

MS. HORWITZ: That's about right.

MEMBER JHA: I understood it was 2.9 percent or something --

MS. HORWITZ: I think Elizabeth has the actual numbers.

MS. DRYE: The 10th percentile is

15.4 and the 90th is 18.2.

MEMBER JHA: So it's a 2.8 percent difference between the best and worst. Now, I mean, we have talked about all the shrinkage issues, but it does concern me, I guess this, the question has been: does it raise any concerns for you, in terms of when I think about this as a consumer, that's a pretty tiny difference, because are you going to make a difference -- are you going to choose a hospital based on 17 versus 16 percent? Probably not. And yet, like a third of the hospitals are in that range.

So is this going to really give us much information is one question, but that's a philosophical question and you may not be able to answer that, but a question that was raised yesterday was: so what proportion of hospitals will statistically be labeled as outliers based on the overall model?

MS. HORWITZ: Right.

MEMBER JHA: That you can tell us.

MS. HORWITZ: I can answer that in a non-answer way, which is, we have not bootstrapped the data yet, so I don't know how many outliers we have.

But I can tell you with some confidence that because we have real volumes here, we are talking about big hospitals, our median of 700-something admissions, we are very confident that even though there's a narrow distribution, that the confidence intervals around any individual point are going to be narrow enough that we will still be able to identify outliers.

The question of whether a narrow distribution means that this is not a quality signal or that it is not useful or, you know, is a separate question, and I would argue, as a clinician, that we are all bad at readmissions and so I think to some extent, this is reflecting just general badness and to some extent it's reflecting the fact that nobody has had any incentive to think about

readmissions yet.

And so primarily it's been driven by utilization and other things. And so I think that as people start to focus more on this as a quality indicator, we will start to see a much bigger spread.

MEMBER JHA: Can I just follow up on that? So your median is around 17 or 16 percent, does that sound right? And the Jencks paper that kind of got all this going, they use all-cause and they were at around 20 percent.

Do you have any explanation for why a four-point gap between the Jencks approach and your approach?

MS. HORWITZ: So the Jencks data is actually totally different data. You will see that the third or fourth most common readmission for -- or condition leading to readmission is psychiatric disorders.

They included psychiatric hospitals and all kinds of other things in

that data that we did not. The second thing is that they use DRG and we use conditions, and the third thing is that this is risk-adjusted and his was raw rates.

(Off mic comment.)

MS. HORWITZ: The mean should come to the same. Right, I think that primarily it's driven by different population.

CHAIR KAPLAN: Laurent, is that up delayed down, or oops. Richard, I think we have time for one or two more questions. Richard?

MEMBER BANKOWITZ: A quick question. With the hierarchical model, you can take into account hospital effects, so you could create a dummy variable for a safety net hospital, and you could discern if there were any marginal differences, and I just wonder if that had been done, and I think it would be useful to do because we are making policy decisions.

So not that we would change what

you have done, but just to understand from a societal perspective if there's a difference.

MS. HORWITZ: So we never put a dummy variable in for safety net hospital. We did this analysis though, whichever it happened to be, looking into proportion of Medicaid as a proxy for that.

Did we ever put that into a model and see how the model changed in performance? We did not do that. But we did see how the performance of hospitals in those quartiles differed.

CHAIR KAPLAN: Thank you. Jim, you have the last question.

MEMBER BELLOWS: Thanks. You made a nice distinction between measures for accountability and measures for improvement and the fact that an unadjusted measure could be used for improvement.

I know in our system, data that really drives improvement is data that we can bring to a local level and stratify and do in

a timely way, and that for us it would be huge to be able to do things like know differences in sub-populations by people who are discharged home versus discharged to SNF, people who are discharged to different kinds of different outpatient settings and so forth.

All those would require information about the difference on the overall performance on the measure the incorporates the risk as well as the raw readmission rate, and all those are kinds of analyses that a person couldn't do with a model that can only be implemented centrally.

So I am wondering how, in the improvement work that you talk about, you would navigate that kind of understanding and provide data that drives, when you can't actually produce the risk-adjusted rate for those different kinds of sub-populations who would let you narrow down in on the problem.

MS. HORWITZ: Yes, well you know, we are outliers for heart failure and for

pneumonia. We can't reproduce that internally but what we do, is we look at raw rates for heart failure.

We have looked at different discharge dispositions and different units and we have broken our data down that way, and we just trust that our raw data from time to time are stable enough, ex our interventions, that we can look at changes and attribute them to our interventions.

In terms of this model, we don't have data that we have confidence in to include things like that, like disposition. We don't believe in the administrative data around that point, so we don't include it in our models.

But people internally look at things like that all the time, and I don't think that they require being able to match that directly to the risk-adjusted measure.

What the risk-adjusted measure does is tell us overall we have a problem, and

then it's up to us to figure out what the problem is and how to handle it.

CHAIR KAPLAN: Thank you very much and I would like to add my gratitude to the committees for your rapid turnaround, rapid cycle productivity at our request.

I think we are on public comment now?

MR. AMIN: Yes. April, can you open the lines if there are any public or member comments to address the Committee?

OPERATOR: Certainly. All lines are now open.

MR. AMIN: Or in the room.

(No response.)

MR. AMIN: Okay. So, no? So we will move then to voting.

CHAIR KAPLAN: Okay, Adeela it's all you.

MS. ADEELA KHAN: Okay, so we know the drill yet, scientific acceptability of measure properties. Are they both reliable

and valid, reliabilities, precise specifications, testing, validities, looking at specifications consistent with evidence, testing, threats to validity, exclusions, risk adjustment, stratification, meaningful differences and comparability in data sources.

MEMBER JHA: Sorry. I am aware that we are voting up and down based on the measure as it is. We have had some discussion about whether there was a possibility to make the vote contingent on any changes.

Is that on the table, or is this a -- assuming that no changes can be made to the measure and the measure is going to stand as-is, are we voting up or down?

CHAIR KAPLAN: I'm going to leave that to NQF.

MR. AMIN: Well, I think there should be a proposal right?

DR. BOSSLEY: I think what would be helpful -- we have done this before -- is to have all of you outline what you think you

would like to see done and then we'd have to hear from the developer on whether that's possible in the time frame we have, and I have looked at Taroon and Alexis in that.

And then you can vote on, based on that contingent, for these changes. We can definitely do that.

Yes, what might make more sense is right now have you vote on it as-is, see where we are with that, and then move to the next.

MEMBER JHA: Got it. So, and if it passes then there's no opportunity to make proposals to modify it, but if it fails then we have the opportunity to --

DR. BOSSLEY: You got it.

MEMBER JHA: -- bring it up.

DR. BOSSLEY: That's it. Yes.

MEMBER JHA: All right. That makes sense.

MEMBER GHINASSI: Let me just go on record as saying I think that's the wrong order. If we are going to seriously consider

-- and I realize that if the decision is to vote on it as-is, then so be it, we should do that.

If in fact the Committee is saying we are open to suggesting modifications, it's my opinion that to vote first does not -- is not in the spirit of that, that this one -- I just want to say.

DR. BOSSLEY: Right. Yes. Maybe it would help -- let's do it this way. We go through this every single time. Shall we see what the modifications are, see if you all are willing to entertain them, and then we will move forward. That's fine.

CHAIR KAPLAN: Okay. If we are going to go down that road what I would really ask you to do is -- we don't, you know, we don't want to be mobilizing a massive effort here to propose a bunch of things to this -- the measures developer, that may or may not -- let's focus on -- I'm going to just do this randomly as -- in my -- I mean you can all

throw you know, stuff at me.

No more than three, okay, so hit your -- get your cards up soon if you, you know, if you've got real issues because I think we are not going to go more than three recommendations.

MR. AMIN: So procedurally this is how we will handle it. You -- members of the Committee can suggest a response, a change to the measure. The developer will respond on whether that's feasible. The timeline that we are considering is that the updates need to be sent to the Committee by the 13th. Remember we are dealing with an expedited review. And the Committee will have to have a call to review these updates on the 15th or 16th.

So, on our current timeline, that's what we are dealing with. Now, whether we can have some flexibility in timeline will be up to leadership at NQF. But this is based on our current timeline here. So --

CHAIR KAPLAN: Okay, and here's

the local timeline. The local timeline is we are over time. So I would really ask people to be very concise and very specific and with the measures developers, very -- it's doable, it's not doable and not a whole lot of, you know, delay. Ashish?

MEMBER JHA: You know where I am going. I would like hospitals' past performance, and personally I would recommend that we use the last five years of performance, as part of the -- as part of what goes into the model for predicting what that hospital's expected rate should be.

MS. HORWITZ: We think that would be a real challenge to get done in a very short time frame, and I would just comment that the downside of that is that it makes it very hard for hospitals that are improving to show improvement.

CHAIR KAPLAN: So that's probably not.

MS. HORWITZ: I think it would be

hard for us to get done in this time frame.

CHAIR KAPLAN: Richard, did you have anything?

MS. DRYE: Sorry, can I just add that we are developing the measure, you know, for and with CMS and so obviously it's not our unilateral decision at Yale too.

So those decisions would have to go back to CMS.

MEMBER BANKOWITZ: Well, I mean I think that approach is arguable. I personally would not support that approach for many reasons, so I don't -- I would not ask the developer to do that.

You know, listening to the discussion, I am mindful of the fact that we want to uncover disparities and so we don't want to bake them into the model.

I just wonder if there's some way, because -- that the developers might suggest a way to at the same time be equitable in our payment. Since we are basing payment on this,

is there a way we can both be equitable and reveal the disparities -- your thoughts on that.

MS. DRYE: So I would just first separate the measure and then how the measure results are used for any payment, and I think we mentioned yesterday that you know, any hospitals that are struggling that have a high proportion of SES or for any other reason, you know, a policy response can be to support those hospitals and that is part of what's in the Affordable Care Act.

But in terms of making a transition on SES, that's, as you know, really challenging and complex issue for -- it's a policy decision and we'd have to -- I can't make any commitment today on whether that would be doable or not, because it's clearly going to be -- that's a decision that CMS ultimately has to make.

CHAIR KAPLAN: So as-is, okay, so that's probably not as well.

DR. BURSTIN: Just one brief comment on that. Again, any payment issues are outside our purview. CMS could choose to, and in fact I think might, you know, perhaps differentially pay based on the patient population. But that's not in the measure. Right?

MEMBER GHINASSI: Yes, thank you for sending the article last night by the way. I just want to read one last bit, two brief suggestions.

Mind set for example, while medical comorbidities may account for a large proportion of risk in some populations. Social determinants may disproportionately influence risk in socioeconomically disadvantaged populations.

Our review found that few models have incorporated such variables. I agree with that point, and I realize the complexities associated with trying to incorporate that into the model.

However, looking for the keys under the light-post, because that's where the light is, isn't necessarily going to help you find the keys.

I would suggest that there be some attempt to bake into the model some measure of that disparity whether it's housing stability over time, employment status, payer mix. There are certain things that can be put into the model which could serve as reasonable proxies for that.

My suggestion is to, rather than simply dismiss them, to include those as a piece. I do think that the article was correct, and my concern is that the primary focus is to improve quality. We may be missing that opportunity.

CHAIR KAPLAN: Can I ask the developers to respond to that quickly. I think you have probably already answered that with respect to Richard, but please --

MS. HORWITZ: So it is certainly

true that socioeconomic factors play an important role in readmissions, and the question is twofold. First, how would we handle that in a model? We could stratify our results on a patient level. We could stratify our results on the hospital level. We could say, we could report separate results for your low-income patients versus your high-income patients, and I am not sure that the public would like to see that. I think that's a slightly peculiar thing to do.

We could stratify our results by hospital, so if you are a hospital that takes care of low-income patients we could compare you only to your peers.

And then we have a funny situation of what if you are better than your peers but you are still worse than the national average, how do we report you? Do we say that you are better? Is that really the message we want to give to consumers? It's kind of a -- it's a challenging thing to think through. We've

thought a lot about this.

And the second thing I would argue one more time, and I feel very strongly about this as a person who is focused on readmissions at my hospital, the impact -- the importance of socioeconomic status on readmission rates is not entirely fixed, and hospitals have a profound role to play in changing that risk and in terms of the way that they perform.

We have cut our heart failure readmission rate by 33 percent in the past year, and we have the same low-income patients that we had a year ago, and we have worked incredibly hard with those patients to really improve their outcomes.

So I am reluctant to endorse the fact that socioeconomic status is immutable and a risk factor that can't be changed.

CHAIR KAPLAN: Thank you very much. We have time for one more. Ashish?

MEMBER LAZAR: Sherrie?

CHAIR KAPLAN: Yes. Eliot?

MEMBER LAZAR: Yes. I don't know where we just left that discussion on SES. I didn't hear the developer say that it was not possible to do.

I would echo the comment before in saying that I'd personally love to see something in there addressing SES down at the patient level, certainly not at the reporting level.

CHAIR KAPLAN: The way I understand this, and help me if I'm wrong here, is that the answer was, the claims data don't -- beyond Medicaid, the claims database, administrative database does not provide you with enough detail to put a credible SES measure into the model. Is that accurate?

MS. DRYE: It would be very hard to define something that people could agree on that we felt comfortable with.

CHAIR KAPLAN: Does that answer your question Eliot?

MEMBER LAZAR: Okay.

CHAIR KAPLAN: Okay.

MEMBER ASPLIN: I have a process question, not a -- I understand our purview is up or down on the measure. Do we ever forward measures with any policy recommendations recognizing it's not -- it's up to CMS to determine how they are going to use them, but do those recommendations from a committee like this survive and go with a measure or not?

Does it just -- because the stratification by SES seems to be the out here, in how the measure is implemented. So you let the variability flow through without putting SES in the model, but you could stratify your sample in groups.

Now, would a recommendation like that survive if it went forward -- the measure got forwarded from NQF to CMS?

DR. BURSTIN: Yes, especially if it's something around stratification. Again, something specific to the measure as opposed

to the payment policy very much so would be in our purview, and that's a very reasonable request if that's something the Committee wanted to consider, that you know, these measures should be stratified.

Again, I don't know how difficult that is to do. I don't know what the implementation issues are. But that would be something I think would be very reasonable to entertain.

MEMBER ASPLIN: That to me is a way that we can address this concern. It's not in the measure development. The measure stays as is. But that's what we would recommend, is that we -- that the use of the measure be stratified according to payer mix and either by quartiles or some other -- maybe just two groups over a certain cut point of Medicaid patients and the payer mix is -- and have two different groups. That would be my recommendation.

MR. AMIN: The Committee can

definitely make that recommendation in the final report. Go ahead Eliot.

CHAIR KAPLAN: Okay Eliot.

MEMBER LAZAR: I'm sorry. I have to strongly disagree with the idea of reporting in a stratified way. I think it -- and again perhaps the public folks, you know the consumer advocates, would like to comment.

But I think it sends a very, very negative message in terms of not having one standard.

CHAIR KAPLAN: Okay so here's where we aren't. I've heard the measures developers really respond to what's possible and what's doable, and what's -- what their limitations are.

So far, correct me if I am wrong Committee, but I haven't heard anybody say -- come forward with a revised recommendation that's -- that the measures developers think they can deliver on. Has anybody heard that?

(Off mic question.)

That's what I'm asking. Is -- because I wanted to see if somebody else heard something I didn't, and if so, because we need to vote, and I'm trying to clarify what we are voting on. So Tanya, are you about to say something?

MEMBER ALTERAS: Well, I just was going to respond to Eliot since he was asked, you know, he said that it was a bad idea. I would say there -- to me there's a distinction, you know, oftentimes we, from my job, we ask that measures be stratified by race, ethnicity, language, gender, but that's really for process measures, because you know, that's where you kind of can identify where the disparity is.

And with outcomes it is a different animal. You know, you don't want to necessarily make those distinctions because you do want to see the same outcomes for everyone. So --

CHAIR KAPLAN: Okay, Brent, tight

--

MEMBER ASPLIN: Can I respond to that?

CHAIR KAPLAN: Yes.

MEMBER ASPLIN: Well my response is that normally I would absolutely agree with Eliot and the perspective that was just raised, when the outcome is really arguably completely within the purview of whatever entity is being held accountable for that outcome.

I just think that there's so much about readmissions that takes place outside of the walls of the hospital, that that's the justification in this case for recommendation that the measure be reported in a stratified fashion.

It doesn't, in my mind, mean that there's two standards of quality. Normally I would absolutely agree with Eliot's comments. I just don't in this circumstance.

CHAIR KAPLAN: So given what we

have heard from the measures developer, because I am now trying to synthesize what I have heard so far, the basic -- the data that you have that most closely approximates SES is Medicaid, and you showed data that there really isn't evidence that there is a distributional difference based on Medicaid stratification.

Would you be wiling to go along with some reporting recommendations stratified on Medicaid?

MS. DRYE: I'm not sure if you are asking would we go along with it, because I think, just, this is a process question, we -- I mean, that's, are you, I think that -- the Committee is considering making that recommendation that the measure would be implemented that way?

I guess I'm trying to --

MR. AMIN: It seems that the question --

MS. DRYE: Do you want a

commitment from CMS in advance that they would implement it that way?

MR. AMIN: It does seem that the question of reporting actually is outside of the scope of what the measure developer would be responsible for.

So really, it seems that that would just be a recommendation the Committee would hold in its draft report up to the CSAC and further on in the process.

DR. BURSTIN: Yes and no. NQF does have criteria that have actually just been updated by our Disparities Committee that are called disparities-sensitive criteria, identifying measures where there are known disparities, and this, I think, one could argue, it's not clear there are known disparities in this are, where stratification is preferred.

And so those measures come out and there's an indication this is a measure that should always be stratified. So I don't think

-- again, I think you need to decide if it's the will of the Committee. We haven't heard that yet.

But I think it could be an accompanying recommendation that goes along with the measure when it goes out for public comment, that suggests the will of the Committee is or is not that that measure should be stratified, or that CMS should look into other ways, I mean, again, it doesn't have to be decided at this moment. It's more of a reporting issue. But I do think it's something we should get a sense of the will of the Committee first.

CHAIR KAPLAN: Okay, so here's the two options now that I hope I am getting right. The two options are one, to go forward with an as-is, no adjustment of hemline, no addition of lace onto the garment, and the other is the accompanying recommendation for reporting if the recommendation for a reporting stratification based on proportion

of Medicaid patients.

So which do we -- which -- is that -- does anybody hear something I didn't hear?

MS. DRYE: Can I just ask also, I don't know how specific you want or need to be, and how you would stratify, quote unquote, for SES, just because, you know, you can do safety net, non-safety net, you can do for example Medicaid, as you know, there's many -- you can do ZIP code, income, blah blah blah, and did you mean by patients within a hospital, or by hospitals with a certain proportion.

So I think those things all actually -- it's been shown by a number of people that those lead to different groups of hospitals and I just -- I just don't know whether you want to get to that level of specificity, the implications of leaving it open, or that -- those are just very different things.

CHAIR KAPLAN: No, we don't want

to be too detailed because I'm afraid that if we put too much specification around it, then we are providing guidance that we really -- we really oughtn't to in terms of our various -- you know, we can get into whole day's worth debate about what the right marker for SES is and how these data should be reported out.

I think a recommendation that they should be reported by stratification would make many committee members more enthusiastic about supporting the measure going forward.

But now I am a little bit confused about procedure. Do we vote on the as-is measure and then the addition, or do we just fold that in?

MR. AMIN: It's overall -- it's just a recommendation on the measure going forward. Is there really --

MS. PACE: But maybe we could just -- we have a simple yes/no question to see how much of the Committee supports that recommendation, I don't know, or maybe there's

a better sense than I have.

MEMBER JHA: Very quickly, so Medicaid is the wrong measure to use because 47, 48 percent of hospitalizations are paid for my Medicare, and so -- and you can imagine that you have lots of poor Medicare patients who are being picked up in these models.

And so there are other approaches -- I don't want to get into what those other approaches are -- but I would only argue that the spread in performance we have seen by proportion of Medicaid really does not capture how much variation there is by safety net status if you use other measures, that safety in the hospital is by proportion of minorities, other stuff, tend to do much, much worse than what we see up there.

So I would -- it's my way of saying there's a bigger problem than -- one that what these data suggest, and there are lots of ways of handling it, none of which we need to get into.

But I would favor that we keep SES on the table as an important issue, and not, in my mind, be affected by those --

CHAIR KAPLAN: Okay, I am going to make then the following recommendation, that we vote with a reporting recommendation attached, to include some reporting stratification by SES and it's a recommendation. It is not changing the core scientific content of the measure.

How many would favor that vote right now, on the measure plus a recommendation for -- no? We are getting the --

(Off mic comment.)

CHAIR KAPLAN: Okay, so that's what we are now -- is everybody clear? Do we have any -- okay. Let me try it one more time. We are voting on the measure as is for its scientific reliability/validity, plus the recommendation that the data be reported with some stratification -- go ahead.

MEMBER LANGBERG: It seems to me we might want to have two votes. The first is as-is, and then regardless of the outcome of the as-is, either -- the up or down, we can then vote whether we want to make a recommendation on the SES issue.

CHAIR KAPLAN: Now are you going to have to do that over again? All right, how many -- let's just do -- let's do a body count. How many people would actually support that recommendation? Hands. Stratification by socioeconomic status, however it gets defined.

DR. BOSSLEY: We have -- we actually have slides ready, in the order that everybody just said. So let me -- Adeela is going to project it in just a second. Let's make sure that you all agree with what we just did.

And I apologize for the messy process.

MEMBER BROOKS: Point of

clarification. When we talk about using the SES or the stratification, whatever it may be for reporting, is that just for Hospital Compare whatever, or how CMS may use it, which I know we can't dictate or tell, in any other way, like for pay for performance, et cetera?

CHAIR KAPLAN: I don't think we can be specific about that recommendation because again, that would take us another day's worth of whatever.

DR. BOSSLEY: Okay, so what we have done, based on the last comment that was said, we can change this again if needed, but we think it's easier to try to do this electronically.

We would first have you vote on scientific acceptability of the measure, because right now that's what you have, and then the second one will be does the Committee support the reporting recommendation and that's a yes/no. Does that seem reasonable?

MEMBER GHINASSI: Just one small

clarification. When you say would report -- vote on the measure, first, I am assuming that means as-is --

DR. BOSSLEY: As-is.

MEMBER GHINASSI: -- without stratification?

DR. BOSSLEY: Correct.

MEMBER GHINASSI: Why would that be different than the vote yesterday?

DR. BOSSLEY: Because you have had new information presented --

MEMBER GHINASSI: New information --

DR. BOSSLEY: -- to you today.

MEMBER GHINASSI: Okay. Thank you.

CHAIR KAPLAN: Okay, Adeela.

MS. ADEELA KHAN: Okay, so on scientific acceptability of measure properties, was the criterion scientific acceptability of the measure properties met. Vote one for yes, two for no.

And you can start now.

MR. AMIN: Okay, I will do a roll call for the members on the phone. Eliot, yes/no, on scientific acceptability?

MEMBER LAZAR: Yes.

MR. AMIN: Patricia McDermott.

MEMBER McDERMOTT: Yes.

MR. AMIN: And Mark Williams.

(No response.)

MR. AMIN: Okay. Thank you.

MS. ADEELA KHAN: I think we are missing one person.

MR. AMIN: Mark, you are -- Mark, are you on the phone?

(No response.)

MS. ADEELA KHAN: We have 11 for -- or actually, 13 for yes, 6 for no.

DR. BOSSLEY: So then the next thing you will vote on is the reporting recommendation.

MS. ADEELA KHAN: Again, one for yes, two for no.

MR. AMIN: Does the Committee support the reporting recommendation, one yes, two no. Eliot?

MEMBER LAZAR: No.

MR. AMIN: And Patricia?

MEMBER McDERMOTT: No.

MS. ADEELA KHAN: We have 8 yes and 11 no. So on usability, to what extent was the criterion usability met, one for high, two moderate, three low, four insufficient.

MR. AMIN: That's the usability criteria. Eliot? It's one high, two moderate --

MEMBER LAZAR: Low.

MR. AMIN: Okay, three, low, insufficient. And -- he said low. And Patricia?

MEMBER McDERMOTT: Moderate.

MS. ADEELA KHAN: Oh, I think I'm missing one person. Wait -- can you press it? Oh, there we go. So we have 1 for high, 8 for moderate, 11 for low, and 4 for insufficient -

- or zero for insufficient.

CHAIR KAPLAN: Okay, so we are voting on feasibility again. Remember that these are all in light of the new information so you are re-voting because of the new information provided by the measures developers.

MS. ADEELA KHAN: To what extent was the criterion feasibility met, one for high, two moderate, three low, four insufficient.

MR. AMIN: This is a feasibility vote. Eliot?

MEMBER LAZAR: High.

MR. AMIN: And Patricia?

MEMBER McDERMOTT: High.

MS. ADEELA KHAN: We're short two votes, if everyone wants to vote again. So, 14 high, 5 moderate, zero for low and zero for insufficient.

CHAIR KAPLAN: Okay, now we do the overall summary vote on the measure.

MS. ADEELA KHAN: So we're asking, does the measure meet all NQF criteria for endorsement, one for yes, two for no.

MR. AMIN: Overall, Eliot?

MEMBER LAZAR: Yes.

MR. AMIN: And Patricia?

MEMBER McDERMOTT: Yes.

MS. ADEELA KHAN: We have 12 yes and 8 no.

CHAIR KAPLAN: I truly want to thank the measures developers again for what you have done to help us, and every single one of these Committee members, for really doing the job and being as diligent as you have been, thank you again very much.

MEMBER LAZAR: Sherrie --

MS. HORWITZ: We thank the Committee as well.

CHAIR KAPLAN: Thank you. Eliot?

MEMBER LAZAR: Sherrie can I make one comment?

CHAIR KAPLAN: Yes.

MEMBER LAZAR: And this really gets back to the issue of SES, and I understand you know, obviously we have not gone with the recommendation about reporting stratification, although I suspect many of us agreed that SES has a very important role in readmissions, and somehow ought to be taken into account.

For me it was simply the issue of reporting by SES and what the -- what my sense of the public perception of the institution would be around doing that.

I would very much like to see, you know, some recommendation statement, assuming, you know, the Committee agrees and it's the will of the Committee, that we do believe SES is important, and that you know, there ought to be some thought given to how to incorporate SES into the model.

That's a little bit more of a general statement than talking about -- or them recommending that something be included

at the reporting level versus the risk level, and I just wonder if other members of the Committee would agree to that, you know, to such a statement, and if NQF feels that you know, that would be appropriate.

CHAIR KAPLAN: Eliot, I'm not going to revisit this at the Committee level, but I am going to ask Helen to comment on drafting a recommendation and circulating it to us afterwards.

DR. BURSTIN: I'd be fine doing that but -- okay go ahead.

MR. KRUMHOLZ: Can I just say one quick thing? This is Harlan Krumholz. One, just to thank the Committee and everyone who has involved here, but your message will be loudly heard.

I just want to let everyone know that the comments you have made, all of the comments, but in addition specifically the comments about the socioeconomic status will go back to CMS. We are taking them and

listening to them very carefully. We understand the importance of this issue. We recognize the care with which you have thought about this issue and are expressing this concern.

We have heard it from others, and what we are challenged to do is to figure out how to manage it, because we on one hand don't want hospitals to be unfairly characterized. We don't want people chasing quality issues that really represent problems that reside within their communities.

On the other hand we don't want to create a -- as you said Eliot -- a two-level system, and we don't want to obscure important disparities that may exist within our society in performance.

So this is -- but we have -- I just want to be clear -- we are listening very carefully to you. We will send this message back to CMS. We are sending it back to our group. And we will work hard to try to ensure

that this is taken under consideration, and we welcome any suggestions from anyone who is listening or anyone who is in the room, and we want to improve.

So it would be great for a statement from the Committee, but the work that you have done has already made an impact and I think that we are going to work to think hard about this and try to figure out how to proceed.

But we welcome any suggestions too, because it's not easy.

CHAIR KAPLAN: Right. I truly appreciate both comments and I -- it's krumholz@yale.edu is that accurate?

MR. KRUMHOLZ: It's harlan.krumholz@yale.edu.

CHAIR KAPLAN: There you go.

MR. KRUMHOLZ: And send them on.

CHAIR KAPLAN: So yes, so I invite anyone to share their recommendations and opinions with --

DR. BURSTIN: And just one more process thing, we will be writing a draft report for all of your review and comment very quickly since it's got to get out tout de suite, but we will try to obviously incorporate the spirit of this discussion in there, even if there's not a specific recommendation.

But clearly this is a major sticking point for the Committee. We would love to have somebody kind of break this logjam and figure out the right approach to allow us to understand the issues of SES but not obscure disparities. So more work to be done.

MR. KRUMHOLZ: And we've listened to all the comments too, SES, but all of them are important to us. So I just want to make sure, it's not like we get approval and we go back and we haven't listened -- we have listened very carefully to all of the comments that have been made.

CHAIR KAPLAN: Thank you again. Okay, I think -- right? We are ready to -- thank you, thank you very much for coming.

MS. HORWITZ: Thank you.

CHAIR KAPLAN: Okay, I'm going to take a poll on what -- Bruce. Bruce, do you have something to say?

MEMBER HALL: I just wanted to lend my support. I hope that our report, that our draft report could just highlight what we have all thought about in terms of threats to the future value here. I want to lend my support to that.

CHAIR KAPLAN: Thank you very much. I'm sure that Helen and the staff of NQF have listened to us very carefully. I would like to do the following, and if I get a massive pushback, please tell me.

I would like to go -- push through this break and invite people to take a bio-break or get coffee or everything on an individual level, and keep moving us forward

so that we -- is there anyone who finds that offensive or would like me to --

MR. SAUNDERS: I would say that unfairly discriminates against the measure developer, who would be talking, and could use one of those breaks.

CHAIR KAPLAN: There is no such thing as a five-minute break.

MR. SAUNDERS: If I could run to the front of the line --

CHAIR KAPLAN: If the measures developer is insisting, then you have got five minutes but tick tock tick tock, I will round you all up and find you individually.

(Whereupon the above-entitled matter went off the record at 9:54 a.m. and resumed at 10:01 a.m.)

CHAIR KAPLAN: Before we get started on the last measures development, the next measures development piece, I would like to invite Tanya -- and really limit it to five minutes total conversation -- Tanya, Michael,

and Leslie to raise issues about the basic accountability/usability issues from each one of their perspectives.

Because I really want to sort of flesh this out in terms of getting it in the transcript and, also, the report by the NQF, if you could give us the succinct issues of concern for this measure, and not more broadly than that, this specific measure for use in the national profile for accountability and quality improvement?

Michael, do you want to go first? And again, please, if you can, keep your comments tight.

On the one that we just passed.

MEMBER LANGBERG: I will be succinct.

The purpose of our work was to give or not give an NQF-approved or validated measure to CMS for its use. We were instructed in the first day, first part of the day yesterday, that the purpose of that use

was twofold. One was accountability, and one was performance improvement.

I am not persuaded -- I am certainly persuaded about the statistical scientific validity discussions we have had, and respect to the vote of the Committee. We haven't had, actually, subsequent conversation again about the usability.

So, I am still stuck over the fact that the reporting will be one to two years after the event. So, we are basically holding facilities for work or experience that may be one to two years old. And the ability to use that information for performance improvement, knowing how rapidly the fields are moving, one to two years later seems to me to be very limited.

So --

CHAIR KAPLAN: Thank you. Go ahead. Go ahead.

MEMBER LANGBERG: Sorry, I lost my thought. If it comes back, I will let you

know.

CHAIR KAPLAN: I'm sorry, I interrupted your train of thought there. Are you sure?

MEMBER LANGBERG: Ah, the only other comment I have made is that I am not persuaded that the hospital-specific metric, in addition perhaps to socioeconomic status and perhaps others that are community-based, adequately assigns accountability to the hospital for the results of the metric.

CHAIR KAPLAN: And that is particular to this specific measure we just passed, right?

MEMBER LANGBERG: Sure. Yes.

CHAIR KAPLAN: Thank you.

Tanya?

MEMBER ALTERAS: I think for the purposes of public reporting that this measure will not be very useful to consumers. And I know the argument has been made the consumers don't necessarily use this information to

begin with, but our philosophy is, you know, we put the best information out there and, hopefully, we will get them more engaged.

And with the advent of an all-condition, all-cause readmission measure, the potential is there for more consumers to use it because the readmission measures that we currently have are not for the types of conditions that people normally go online to check their hospitals for, if they are having a heart attack.

But when you are broadening it to all conditions, we really did see this great potential there. So, I am not really going to speak to the accountability issue because I think there's a whole lot of issues there in terms of how well this will work for holding hospitals accountable.

But in terms of transparency and public reporting, I didn't really see -- and I apologize that I missed some of this morning's discussion -- I still am skeptical

that the results, when reported on Hospital Compare or on other websites, will provide consumers with useful information that they can make choices on.

CHAIR KAPLAN: Thanks very much.

And Leslie?

MEMBER KELLY HALL: My concern on usability is really more recommendations on how the reporting is used and explained well for the public. In rural communities like Idaho, if this information is proved useful to consumers and they seek it out, can a patient who is not reviewing this information with a high degree of understanding make inappropriate self-selections?

And the distance factor for hospitals, the lack of community safety net or connectivity of others to provide social support, to provide followup, is even heightened in rural communities where you might have 200 miles between hospitals.

So, I just caution the group to

take a look, when reporting specifically around rural hospitals, and how do we provide that kind of information that is useful?

CHAIR KAPLAN: Thanks very much for those comments. I am sure that Helen will make sure that they are in the report, the Committee's report, as reflected concerns along with lines of accountability, quality improvement, and reporting, as I heard it. Thank you very much.

So, should we tee up the --

MEMBER GREENWALD: Could I just make one other additional question about that one, just as a process piece, very briefly?

I respect the opinions of my colleagues here. The question I have about that, though, is, how were those comments specific to this measure? I think many of us have a healthy skepticism about the utility of an all-cause readmissions measure, generically speaking, in terms of its utility and usefulness at the consumer or hospital level.

I am not sure how that reflects this specific measure individually.

CHAIR KAPLAN: Right. I am going to invite those with a consumer perspective in general to raise their concerns focused on the specific measure and give them to Helen for the report, because I think that will allow them a little more time than we have here to really develop their issues.

MEMBER ALTERAS: In 20 seconds, from the consumer perspective, the issue isn't with an all-condition, all-cause readmission. To us, that is a glide path to system redesign. So, it is really more about the public reporting and the transparency and how it is reported.

CHAIR KAPLAN: Thank you.

We have the next measures developer. We have 10 minutes allotted. If you could make your presentation succinct? And ready, set, go.

MR. SAUNDERS: That would be a

first.

(Laughter.)

So, thank you again for the opportunity to come back here to talk about this.

So, while these were generated new this morning to put into the presentation here, we had done these kinds of graphs before in the previous model years. We just hadn't done for the current model years.

But these are the plots of the actual versus expecteds for the risk deciles. We have additional sorts of error here. It is my ability to draw a red line at a 45-degree angle onto the graph. But you can see, if you were to assume that that red line were going to the origin, trying to map it out there, that our blue dots are lining up like they are supposed to on the red line.

If we can scroll down -- sorry, we want to scroll up here.

So, this is the Medicare ages, 18

to 64. This is our Medicare 65 and older. I think my line is tilted a little bit here. The bottom should be above a little bit, and you will see that they are in line.

If we scroll up again, a similar pattern and a similar drawing error on the red line.

But if we scroll up, or sorry, if we go to the tab above, we can see what the actual differences are here, that we are talking less than 1 percent at each of the deciles.

To Laurent's comments, this is in the predicted, in the model dataset. So, it is going to be naturally better. We didn't have a dataset to validate on this morning. But, hopefully, this is responsive to your concerns about that. So, we feel like we have adequate discriminate ability in the models.

Let's see, our second question I believe we were to respond to is the SES issue. So, I was chatting about this with

other folks at NCQA who have been in this business longer than me and said, "You big dummy, there's plenty of examples of high SES variability between plans within the same markets." And so, I retract my overstatement and the Redd Foxx references.

I think it doesn't change our opinion that we feel, similarly to Yale, that risk-adjusting out variability due to SES is not appropriate, but we also have heard the feedback of the group that this is important.

We feel that we take this into account to some degree through our measurement through separate product lines. So, we measure the commercial groups separately, the Medicare group separately. If we were to build this measure for Medicaid, we would have a separate measurement for them, and would hope that there is sufficient similarity within those groups.

I think, apart from the concerns that were raised about the difficulty of

actually measuring SES accurately, feasibility is one of the four NQF criteria in this measure. That translates to us in terms to the implementability with the data collectors and the people that are reporting the data.

For us, collecting the data from health plans would require the health plans to go beyond their normal data collection processes to collect their readmission rates based on individual patient zip codes. We would have to come up with a way to link, then, Census zip code information, assuming that we even believed those numbers as representative of SES, into that dataset, to then do the adjustments.

So, we think that in terms of both kind of the conceptual rationale for the inclusion, which I realize the panel disagrees with, but also the practical burden of that feasibility is an important criterion for the implementation here, that we would argue that it would be necessary to leave SES out of

measurement. But we would certainly be open to exploring ways to shoehorn it in, if there are ideas and suggestions.

I think in terms of the usability of the measures, I think we have sort of talked about this as the Yale measure is a CMS measure. The NCQA measure is a CMS measure as well.

CMS has financed the development of this through our performance measurement contracts for our Geriatric Measurement Advisory Panel. And so, we have built a whole suite of measures for health plan monitoring, for Medicare Advantage plans.

So, I think, first cut, we think that the measures are usable and are important because CMS has thought that they are important. It is important to have the accountability both at the hospital level and at the health plan level, and they have recognized that, even if they are funding it through different groups within CMS.

We think that there is opportunity for improvement in addressing the problem of readmission from both sides, whether from the hospital perspective or from the health plan perspective.

As a second element to that, I think we are in similar line with how Leora had sort of described how they use their performance measures with her hospital at Yale. We think of this as the same process for our health plans. We provide a national standard and benchmark for how to measure things that would allow for consistent comparison across plans, but it is also useful -- and so, that gets at the accountability function -- but is also usable at the quality improvement level as a benchmark to guide your rapid-cycle improvements. No one, I don't think, will ever have rapid-cycle versions of this measure in doing the risk adjustment. There will always be that lag.

But we think that for hospitals who are able to look at their real-time results on readmissions, or for health plans to look at their real-time results on readmission behavior, and to link that to the performance on these metrics, HEDIS is a large national program, and health plans invest a significant amount of money to improve those rates. Central to that has been having our metrics as a guide for their individual quality improvement efforts.

So, we think that, based upon our experience, that our readmission measures would be in the same vein and have the same utility, and has been the voice of the health plans that have responded in our public comment.

MR. AMIN: Robert, we are at 10 minutes.

MR. SAUNDERS: Okay.

MR. AMIN: I know, if you have a few more things, feel free to take another

five, but --

MR. SAUNDERS: No. I think that the last thing that I would say is that, well, I think the last thing I just would want to mention, because I had been thinking that we were going to talk just about the acceptability, the scientific acceptability, just slide to the feasibility to the component.

We would say that the primary evidence of the feasibility of our measure is that we have already collected it for one year; that the measure has been implemented by health plans. CMS is already in the process of using the measures within the star system, that it is both for health plan choice and for incentive processes.

NCQA has every intention of using the measure and the results for its own public reporting processes and products, like Quality Compass. And we have the opportunity to include it in other things like our health

plan rankings and other types of things.

So, this measure is very real, very feasible, and very usable by a variety of folks, and has been in demand by all perspectives, provider and plan and consumer and employer group, and everyone else.

We think that there is utility to having this health plan perspective in combination and in complement to harmonize with the hospital-based measure.

CHAIR KAPLAN: Thank you very much.

Now this measure is open to the panel for discussion. Questions?

Ashish?

MEMBER JHA: Just a point of clarification, and thank you as well for all the terrific work. It is really helpful.

The risk-adjustment model that you guys use -- forget the whole shrinkage issue that we talked about and we have belabored -- just the strict risk adjustment, my sense from

looking at the data from what you guys presented is it is very, very close to the CMS risk adjustment in terms of the approach, the HCCs, what goes into the straight model.

Can you tell me whether that is true, how close it is, if there are important conceptual differences? Again, we are not going to talk about the hierarchical modeling part of it, but just straight risk adjustment.

MR. SAUNDERS: Sure.

MEMBER JHA: Are there important differences?

MR. SAUNDERS: If I could ask, Alexis, if you could skip to the desktop, there is one other sheet that we had put in, the NQF harmonization sheet in the middle there.

So, we kind of went through looking side-by-side, the exclusions between the two models -- and I haven't verified this with the Yale folks; they can obviously correct me where I am wrong -- and looking

side-by-side, also, on the risk-adjustment strategy, I think for the most part, if we are looking at -- we have obvious difference on accountable entity, but we have similarities on the continuous enrollment criteria, similarities in age. We are looking at acute hospitalizations. We are handling transfer similarly; we want the last place.

If we scroll down -- I don't believe they have a restriction to have an overnight stay for the intense hospital stay, but we have you have to be overnight. So, we get around that sort of observation room problem that people were concerned about yesterday.

So, I think side-by-side things are pretty similar. We include the behavioral health; they don't. But, for the most part, those types of exclusions about the definition of the denominator and the numerators is pretty consistent.

If we go to the risk-adjustment

strategy, sort of doing a hasty sort of comparison here between the models here, obviously, they are doing the hierarchical; we are doing the indirect standardization through the logistic model, for the reasons described yesterday. The clustering problem is kind of systematic, and there is not very much that we can do about that.

In terms of the risk adjusters, they are using age; we are using age. They have surgery built into their model; we have it as a covariate. They have their index condition. They are using the CCS categories; we are using the CCs from the CMS-HCC system. But there is probably substantial overlap in what those categories are.

And then, the comorbidities, I believe looked like they were using the CCs, and we aggregated those to the HCCS. But, conceptually, pretty darn similar.

I would sort of yield to the Yale folks, if there is any misstated here.

Oh, yes, I'm sorry. So, I think I did have that on the previous tab.

So, they do have the exclusion for planned readmissions. I think another difference is that they count readmissions as, allow them to be an index event. When we were developing this, we were modeling after the condition-specific measures. And so, we kind of followed that logic. That is a rationale behind that. There is also a rationale for going the other direction.

But I think these are pretty similar processes were looking at. Adjustments for patient demographic attributes that are available, your condition that you are in the hospital for, and what your comorbid conditions are. I think, to that extent, they are pretty darn similar.

CHAIR KAPLAN: Thank you.

Bruce?

MEMBER HALL: I would add that I don't know if you intended to reflect it on

the previous, but I would add that the structure of cohorts, of infinite cohort models is a difference between your approaches.

So, I took to heart what Sherrie said yesterday, that there is no reason necessarily that we wouldn't expect a plan measure and a hospital measure to need to be different. On the other hand, I am not seeing any axis here where there is a convincing pressing need for these models to be different.

And so, I would ask you to respond to the notion that we basically would be putting forward a system where we would have to do calculations one way at the plan level; we would have to do calculations a different way at the hospital level. I think that would be an unnecessary burden. I don't see any pressing reason for the differences in these decisions.

I think your measure would work

with their decisions, or vice versa. I am not saying it should go one direction or the other.

But, again, Sherrie's point was there could be a good reason to need a difference in modeling, but I am not seeing that. And so, I wonder if you could comment on it.

MR. SAUNDERS: Sure. I think it comes to the definition of harmonization. Does harmonization mean identity? And so, at least at one level, there is a harmonization in terms of the model that we have age; we have handling surgery. We are handling index conditions. We are handling comorbid conditions.

If there is a need to go to a lower level, I think we are certainly open to considering looking at that. And I am sure that they would be open to considering doing that as well.

And I think what the harmonization

process is is figuring out what makes sense to implement. I mean, from our perspective, we have practical issues that argue for us to continue using the HCCs, and that is sort of built into what our index conditions are and what our comorbid conditions are. We use the HCC risk-adjustment process for our resource-use measures.

And so, it is something that the plans are familiar with. They have been doing that for, this is their fifth year of collection on that. This will be their second year of implementation with this measure.

So, we are out in the field, and we have a body of experience working with this. And so, we have an institutional pressure to avoid that, but it doesn't have to be super-binding.

But, for the practical considerations, as the measure developers, and to do this in the business of accrediting healthcare organizations and standardization,

there is a real burden to adjusting the processes from what are in place.

So, I think we would argue that it would be judicious to take a level of harmonization that seems appropriate to the purposes. We would think that there is alignment in the comorbid conditions, and there is alignment in the index conditions, and that the numbers and the namings of the things are not going to be all that different, and there will end up being some sort of collapsing and negotiation process for linking those.

But I think, for the same conditions, then we are going to get the same results.

MEMBER HALL: I actually agree probably very, very strongly that you will get very similar results, which just leads me to, again, push the notion that part of harmonization is about simplicity and burden, and the simplicity and burden for the

hospitals and the simplicity and burden of usability for the patients who are going to be told, well, that measure is calculated a little bit differently.

MR. AMIN: Bruce, can I just jump in here real quick, procedurally?

We won't really go into the discussion about harmonization until this measure is actually recommended for endorsement. We have some segment of the agenda to discuss that. But let's really have the comparison because, ideally, we would have Yale being able to respond to concerns on both sides.

MEMBER HALL: Okay.

MR. AMIN: So, let's really just sort of --

MEMBER HALL: Okay, fair enough.

So, the second part of my question is, I wonder if you could again comment on what the significance is of this measure being listed as proprietary. It is listed as

proprietary. So, does that mean there are any issues with respect to having this in the public sphere?

MR. SAUNDERS: No.

MEMBER HALL: Presumably, CMS would have, if it wanted to implement it --

MR. SAUNDERS: No, it goes into the public -- yes.

MEMBER HALL: -- they would have to make a call for any vendor that could do this. Is that correct?

MR. SAUNDERS: So, there may be sort of an interpretation issue of our understanding of how to answer the question. It is proprietary in the sense that NCQA has a copyright on the measure, but, like every other NQF -- we have several dozen NQF-endorsed measures, and they are not only proprietary, but they are all shared out and they are publicly shared and available. That was part of our rationale for using the HCCs, is that it is a freely-available thing. So,

we don't have to get into the issues of some of the risk adjustments and those types of things, or things with DRGs and the non-standardization of that.

DR. BURSTIN: So, just one final clarification. There is proprietary, meaning it is their intellectual property, but it is without fees. So, just to be clear.

CHAIR KAPLAN: Thank you.

We are going to go Jim, Paula, Richard, and then Jeff.

MEMBER BELLOWS: Thank you.

I wanted to do two things. One, speak to the usability and also ask a specific technical question.

On usability, you sort of deferred to the same explanation as we had for the Yale measure about accepting a difference between measures for accountability and measures for improvement. And I hypothesized that nobody would be able to bring those together.

For our system, where we do have

both plan and hospital data, it is huge that we can implement this on our own independently on a per-member basis. And we do provide this in near real-time and do use a version of your measure for performance improvement and are able to report it on both a plan basis and a hospital basis and a discharge clinic basis, and all those things.

Maybe in everybody else's system they can hold accountability and improvement in two separate parts of their brain, but in our system there is a tremendous desire to bring those together. The fact that this measure allows us to bring those together gives it just phenomenal properties for usability that other versions don't. So, I wouldn't let go of that.

I do have just a fairly-narrow question, which is that I see up there it says the reporting metric is the risk-adjusted readmission rate, the O-to-E ratio multiplied by the national average. And I understand in

principle that that is the right treatment.

When I looked in the documentation, I couldn't actually find that that is the metric. What is really in the documentation is that the reported metrics are the individual pieces of that, the raw rate, the average adjusted probability, the stratification. There are some little details that are required to produce that risk-adjusted readmission rate that are actually I think not part of the submission, including the exact way that you are aggregating across the age brackets and the exact values that you are using for the national average observed rate, and so forth.

So, am I right that actually saying that the reported measure is the risk-adjusted readmission rate on the slide is different than what is in the submission? And if so, are you making all the details for making that final calculation available to the public?

CHAIR KAPLAN: I would just ask, in the interest of time, make sure we get to everybody, if you could get your answer to be concise, we would be very grateful.

MR. SAUNDERS: So, because we haven't publicly-reported this measure ourselves yet, we have been sort of working out the details. And so, the risk-energized readmission rate is a measure we want to do. We have not implemented that for public reporting yet. And so, we would work out those details.

MR. AMIN: Just for a clarification point, if there is a difference, I think we need to know which one is the actual one that is being voted on for the Committee. So, if there is really a difference in the way that it is being reported, it is pretty clear that we know which one it is.

MR. SAUNDERS: It is what is on the screen.

CHAIR KAPLAN: Okay. Thank you.

Paula?

MEMBER FOLTZ: Yes, I do have a concern that planned admissions and, well, basically surgeries and rehabs are not excluded. Basically, from a plan perspective, I would think that that would be an efficiency measure for them.

Because coming from a hospital with Level 1 trauma and burns, we do an awful lot of staged surgeries and send people to SNFs, and then bring them back when they are eligible for rehab. So, I wouldn't want us to keep people in the hospital until these things -- I think there's a lot of efficiencies in excluding those two populations.

MR. SAUNDERS: I believe we have exclusions for the rehabilitation population. We are looking at acute hospitalizations, not rehabilitations.

In terms of the planned admissions, I think this is where we get to

end up needing a Bonferroni correction for advisory panels. We went through our process, through our Geriatric Measurement Advisory Panel and our Committee on Performance Measurement. And now, we are to the NQF Steering Committee here.

We initially had a set of exclusions for conditions that are likely to have planned hospitalizations. So, the active cancer treatment, we would have made a person-based exclusion for that. Organ transplants, things that are sort of likely to have planned hospitalizations or planned readmissions.

Our Geriatric Advisory Panel was in favor of that. Our Committee on Performance Measurement was against that. I think the rationale is that to be accountable for all the hospitalizations and that we are going to measure the plannedness of hospitalizations with error either way. Rather than introduce an error of unknown magnitude in our definition of "plannedness",

to let the risk adjustment handle it.

That is not to say that we can't revisit that. Certainly, if we got to the harmonization stage, we would be open to exploring the handling of planned admissions. So, certainly, the Yale example is a fine method for identifying that.

CHAIR KAPLAN: Richard?

MEMBER BANKOWITZ: A specific question and then a general. The specific question: how does the HCC method incorporate the present-on-admission flag, if it does? I am wondering if you separate out comorbidities from things that occurred in the hospital, like an iatrogenic renal failure, would that be treated the same way as a patient who came in with renal failure?

MR. SAUNDERS: I don't know how it handles present-on-admission. That is something I would ask our coding panel or our coding experts at NCQA. But I could reasonably find out.

MEMBER BANKOWITZ: Can I ask you if the Yale method addresses complications and comorbidities? They are two different things.

MS. HORWITZ: So, two questions. First, present-on-admission coding is not yet reliable. And so, we don't use it. But as soon as it becomes reliable, we would be glad to use it.

And so, in the meantime, what we do is, if there is a comorbidity that is only present on the index admission that we think could conceivably also be a complication, for example, renal failure, we do not count it. But if that comorbidity is present on previous admissions or if that comorbidity is something that we don't think would logically be a complication, then we do count it.

MR. AMIN: Sorry. Just, again, as a procedural matter, it is extremely important that we focus this discussion purely on the NCQA measure and its own merits.

MEMBER BANKOWITZ: Right.

MR. AMIN: I know that there are references back and forth to the Yale developer, but really it is inappropriate. We really should keep this at NCQA at this moment. We will discuss selections around harmonization if this actually passes all the way through.

MEMBER BANKOWITZ: And this is my general issue: the issue of harmonization really impacts usability. I am concerned from the provider point of view. I am a large hospital. I have got one, let's say, predominant commercial plan. And the commercial plan tells me, based upon one methodology there is a problem, and CMS tells me, based on its methodology, "You're looking good." This drives the hospitals crazy. This introduces chaos.

And so, if it is not harmonized, it is not usable.

CHAIR KAPLAN: But we are way ahead of ourselves yet.

MEMBER BANKOWITZ: So, I can't even vote on whether it is usable unless I know what is coming down the road in terms of harmonization because, if it is not harmonized, it is not usable, in my opinion.

DR. BURSTIN: I sort of understand where Richard is going. I think this is an issue. The question would be, if you are going to vote on this measure, are there issues with this measure as you see it? That would potentially be something you would want to address.

Obviously, it is hard to completely disentangle what we saw earlier versus this. But I think, for example, just to put one on the table, I think the issue of the lack of exclusion of planned readmissions is one that I think the Committee should certainly consider. It is in ACA, clearly saying -- you know, again, this isn't subject to ACA -- but I think, just again to Richard's point, it would be very difficult for

hospitals to get both measures, one excluding the planned and one not getting excluding planned.

PARTICIPANT: Explain what ACA is.

DR. BURSTIN: I'm sorry, the Affordable Care Act. So, the health reform legislation specifically indicates that they would prefer they want readmission measures with planned readmissions excluded. Correct, Nancy, my ACA person?

MS. FOSTER: They wanted unrelated such as planned. So, all unrelated such as those that are planned.

CHAIR KAPLAN: So, other questions?

Bruce?

MEMBER HALL: So, technical/procedural, for Helen or Taroon, or whomever. So, if the will of this Committee is that these two measures, if both approved, would potentially be competing, what would our obligation then be to call one versus the

other? Is that not a relevant issue? Is that how we get over saying, well, we got both of these through, but then do we have another decision where we say, now that they are both through, we can only really endorse one?

DR. BURSTIN: I think that is really the discussion, the issue you are going to have to ponder. I mean, the question will be, assuming these both go through, can you justify having both?

We specifically asked the Yale/CMS folks last night, can their measure be used to the plan level? The answer was, "Not yet, but maybe."

We don't have a health plan option on the table. We have heard from the Committee that having a health plan measure, particularly for purchasers and employers, is particularly useful, and probably large plans, what we just heard from Jim.

So, I think the question you are going to decide is, first of all, do they both

meet the criteria? Are there harmonization issues? For example, just to throw out that one example of excluding planned readmissions that you might want to grapple with.

And then, I think the ultimate question is, are these complementary measures or are they competing measures? That, I think, I something we will work through to follow.

CHAIR KAPLAN: I think, though, for me -- and correct me if I am wrong, you guys at NQF -- we have got to stay focused on whether this meets the scientific rigor that we need it to pass before we have the issue of harmonization and putting everything in the blender and coming out with something good.

MEMBER HALL: I guess I was just wondering whether we can put it forward. But, then, the harmonization can be a roadblock at that second point. But I think the answer is no.

CHAIR KAPLAN: So, we are going to

have Jeff. And then were you going to put your -- are you sure? Okay.

MEMBER GREENWALD: So, I am curious methodologically about a decision you guys made that isn't wrong; it is just a choice. And I am wondering if you can explain how you came to that decision, which is the choice to exclude hospitalizations within 30 days. That is to say, the readmission can't be an index admission within the next 30 days.

My sense of that is that lots of studies do exactly that and various models have done exactly that. It has always struck me as less-patient-centered approach and fairly problematic, in my opinion. Because of that, it has its role, but I think probably not here. And I am wondering what your thoughts are.

MR. SAUNDERS: So, we have seen it done both ways. We are not ideologues about this in any way.

I would say that our defense for

how we did this is that, first of all, these are not independent events. And so, we feel like, for an accountability purpose, the mistake was made at the first readmission and that these subsequent readmissions, we have already penalized you on that process.

Now you can certainly have greater magnitude and have additional hospitalizations and turn it into a "how many" measure. And I think that that gets at a different aspect of care.

I think, let's see, it doesn't have to be perfect here. We made this decision initially that, for accountability purposes, that one of that triggered was enough to penalize, to hold accountable. And sort of, statistically, we didn't want to have the impact of the repeated readmissions and the repeated impact of events influence the scientific validity of the model parameters.

MEMBER GREENWALD: Yes, if I might just follow up, I am not sure I agree with

that approach. The not significant minority of patients who have multiple readmissions within that time period of concern represent a significant burden, both financially as well as resource utilization. They also represent opportunities for interventions to be either repeated or initiated.

And given that the highest respecter for readmission is prior admission, it has always struck me as an odd approach to ignore that repeat-offender cycle as an opportunity to really sort of influence how you approach the process.

MR. SAUNDERS: And --

CHAIR KAPLAN: Excuse me. I am sorry to cut this off, but if we don't get to one more question, we are short the discussion. So, I am going to give you like really a five-second response to that issue, if you can. And otherwise, we will go on to Laurent's question.

Can you give a really concise

response to Jeff's issue?

MR. SAUNDERS: No, that's fine. I understand.

CHAIR KAPLAN: Laurent?

MEMBER GLANCE: I just want to get back briefly to Richard's comment. I think it is very important in terms of the scientific validity of this methodology.

In order to be able to really determine whether this is a reasonable model, I think we really do need information on how you go about distinguishing between comorbidities and complications. Because, to a large extent, the validity of the model hinges on being able to make that distinction. If you credit complications as if they were preexisting conditions, that has very, very important implications for the adequacy of risk adjustment.

MR. SAUNDERS: So, our comorbidities look back at the prior 12 months and the discharges that are on the -- I

believe it includes the index hospital stay.

MEMBER GLANCE: So, you do have a look-back period of 12 months, too?

MR. SAUNDERS: Absolutely. Yes. So, if that wasn't clear, our comorbid conditions is a 12-month look-back, looking at inpatient hospitalization records, at hospital outpatient records, professional services records. We require face-to-face visits for the service, so we don't have rule-out diagnoses being triggered on sort of lab or imaging types of things. So, we are truly getting the comorbid conditions.

CHAIR KAPLAN: Thank you very much.

I am just going to clarify one thing from yesterday that I was still not certain about. But if the plan contracts with one, and only one, hospital -- let's just make up that scenario -- the plan is the hospital in terms of precision of estimates. You can't distinguish the two.

If the plan contracts with multiple hospitals, and it is hospital readmission rate, the assumption is that the attribution is for readmissions at the plan level, but you don't know how much precision you have in that estimate or how many hospitals it takes to get a precise estimate of the plan's performance of readmission.

You don't have data on that right now, right?

MR. SAUNDERS: We do not.

CHAIR KAPLAN: Okay.

MEMBER KELLY HALL: Sorry, one question. The plan has the ability to select where they do business. And we talked a lot about socioeconomic issues at a hospital level, but plans have the ability to say, "I want to go after Google because my average age is 32" versus "I want to go after the local bus-driving union where the average age is 57."

And so, how do you accommodate

that sort of self-selection as plans seek business when they do the comparison initially?

MR. SAUNDERS: So, I think this will be a problem for every measure in HEDIS, that every health plan is going to face this incentive for every single measure that we have. And it has not been a barrier to endorsement of any other measure.

MEMBER KELLY HALL: Are any other measures so tied to hospital-specific activities as well as community activities that the hospitals don't have control over?

MR. SAUNDERS: I would wager -- I don't know the entire HEDIS side; Helen might know of a few -- that they definitely hinge upon community-based measures. There are definitely community-based measurement ones. I am not sure that we have any specific to hospital and the separation effect, but definitely the community-based treatment.

MEMBER KELLY HALL: So, self-

selection is an issue for the plan?

MR. SAUNDERS: I think self-selection, plans are always competing on these things, and it is endemic to the process. I think it would be unfair to expect any risk-adjustment model for any particular measure to resolve all of the market selection factors on either the provider side or the health plan side.

CHAIR KAPLAN: Okay. We are going to take Ashish, and then that is really it. And I would ask both of you to be concise, concise question, concise response.

MEMBER JHA: So, this is just a quick comment. It struck me, as I have been listening to this, that both the challenges we are bringing up for the NCQA model, most of those issues are nearly identical at the hospital level.

And all the harmonization issues, which I think Richard brought up, are critically important, but at least my best

read of it is there are no huge challenges here that you can't get over, if both sides are willing to work and we do that. So, I think we can get there.

And I guess one quick, last comment is one of the things that we have all brought up over the last day and a half has been how accountable at the end of the day is the hospital, when we thought about the hospital for readmission. What I have always found really attractive about the health plan thing is you do get this opportunity to look beyond the hospital at a broader set of activities.

And so, in terms of places and people who can actually make a difference -- and whether they will or not is a different issue -- it strikes me that there is a distinct strength of the NCQA health plan approach which make readmissions, at least to me, much more palatable a quality measure than when we were --

CHAIR KAPLAN: So, Ashish, what is your question?

MEMBER JHA: No, just a comment. No comments?

CHAIR KAPLAN: Okay. We are not going to harmonization yet.

MEMBER JHA: No, no, no, no.

CHAIR KAPLAN: Okay.

MEMBER JHA: Right. I was just saying I think the harmonization issues are all solvable.

CHAIR KAPLAN: Right, but that is not right now what is in front of us. What is in front of us is, is this measure reliable and valid?

MEMBER JHA: Fair enough.

CHAIR KAPLAN: Okay.

MEMBER JHA: I guess the reason I went down this road, Sherrie, is that I think brought up by at least several people is concerns about harmonization might affect how people think about these issues. And I was

just trying to make the point that I think those are all solvable if we go down this road.

CHAIR KAPLAN: Yes.

MEMBER JHA: That's all.

CHAIR KAPLAN: That is an "if", and everybody should keep that in mind. So, as we go forward, this is, is this reliable and valid for estimating the hospital readmissions of plans? Got it, everybody? So, we are focused on, is this a reliable and valid measure of hospital readmissions of plans, and do we have enough information in front of us to make that decision?

So, now we are at the point of inviting public comment.

MR. AMIN: April, can you open the lines?

And anybody in the audience here, anybody want to make a comment or address the Committee?

OPERATOR: Absolutely. One

moment.

MR. AMIN: Please.

OPERATOR: All lines are open.

MR. AMIN: Nancy Foster?

MS. FOSTER: A question of process, Madam Chair. This comment would not be related specifically to this measure, but I would like to make it sometime during the course of the day. Can I make it now or later, your choice?

CHAIR KAPLAN: I think for the purposes of this group, Nancy, unless it specifically relates to this measure, I would like to keep it clean. But if and when we go forward, or if this relates to this measure specifically, great. If not, I am going to invite you to do that during the next period of discussion. Is that okay? Or does it relate to this?

MS. FOSTER: It does not relate specifically to this measure, but I would request to be able to do it before you break

for lunch because I have to leave.

CHAIR KAPLAN: Yes.

MS. FOSTER: Thank you.

CHAIR KAPLAN: We may not have lunch.

(Laughter.)

So, are there any other issues of pressing concern that would relate to this specific measure that would help you to make a decision about the scientific reliability and validity of this measure?

The usability is going to come up, but not until -- I think to be fair to the other measure developers, we should do it, the process, the same way, and then invite those with sort of the consumer perspective to comment in the way we just did for the last measure, just to make equity.

Okay. So, Adeela, are we good to go?

MR. AMIN: Well, before we get there real quick, as part of the procedural,

is there any recommended modifications to this measure, based on the new information that was presented? It doesn't appear so, but I want to offer that as parity.

MEMBER GHINASSI: Again, procedural point. There was a recommendation made about adjusting for factors outside of the socioeconomic status, whatever the final label was on that. I don't see any reason why this shouldn't also have that accompanying it, unless I am missing something here.

CHAIR KAPLAN: That wasn't part of the measure vote itself. It was a recommendation back to NQF.

MEMBER GHINASSI: No, I'm aware, and I guess procedural. I don't personally see any difference in the predictive validity of this compared to the other measure. And since the group spent considerable time, and the Committee was generous enough to acknowledge that, I just want to make sure that we are not failing to include that for

this, as we did the other one.

MEMBER ALTERAS: This was raised before, but I would just pull out the issue of excluding planned readmissions and looking into that more carefully.

MR. AMIN: Is that a recommended modification or a condition? Or is it just something to make note of in the report?

MEMBER ALTERAS: For me, I would say a recommended modification.

CHAIR KAPLAN: So, how many, a straw vote, how many believe that that needs to be included in what you are actually voting on right now?

Specify it again, Tanya, what you want included.

MEMBER ALTERAS: Well, this is a recommendation that NCQA look into excluding planned readmissions from the measure.

CHAIR KAPLAN: So, I will sharpen it up. NCQA should move forward with only unplanned readmissions. Is that the

recommendation? Okay.

So, how many think --

MR. SAUNDERS: We are happy to do that.

CHAIR KAPLAN: So, how many thing that should be added to what we are voting on?

(Show of hands.)

Okay. Can I confirm with the developers that that is doable?

MR. SAUNDERS: It is.

MR. AMIN: And the same time restraints that we discussed before. I can read them off. All those modifications would need to be available to the Committee by the 13th of December.

Adeela, do you have that vote, voting slide?

MEMBER BELLOWS: It is almost impossible to exclude just all unplanned readmissions, no matter what technique you use. If we could just modify it a little bit, so they are not forced into failing, that

would be good.

CHAIR KAPLAN: No, it is planned readmissions we are excluding, not unplanned, right? Planned.

MEMBER BELLOWS: Right, but maybe I phrased it wrong. But there is no methodology anyone has developed to exclude all planned readmissions.

MR. SAUNDERS: As an analogy, that plannedness is a concept that is measured with error, and we will get as much as we can.

CHAIR KAPLAN: As a measurement scientist, everything includes error, just so you know.

(Laughter.)

Okay. Are we ready to vote? No.

MEMBER McDERMOTT: Can there be a second recommendation explored?

CHAIR KAPLAN: Okay.

MEMBER McDERMOTT: Related to the group readmission factor of a case after it has been identified as a readmission, the

sequential readmission. There was good discussion about the error associated with that and the bias associated with that. I would bring that as a second modification, that they would remove that stipulation from this measure.

CHAIR KAPLAN: So, the question is whether repeated readmissions, the way I understand the issue, are not counted as readmission; it is part of the same profile. So, it is a dichotomous variable as opposed to the readmission then counts as the primary admission, and then, subsequent admissions are readmissions for the first readmission.

MR. SAUNDERS: That is correct. It is treated as an episode of hospitalizations.

MEMBER GREENWALD: So, the question is, would we change that from allowing a readmission to serve as an index admission for subsequent readmissions. And I would support that as well.

CHAIR KAPLAN: Are you comfortable with that, Measure Developer?

MR. SAUNDERS: We believe that we could deliver a specification by December 13th that included that.

CHAIR KAPLAN: All right. Now we have a bunch of confusion in the group.

So, Ashish, do you want to try to articulate what the confusion is?

MEMBER JHA: For me, I would argue that -- and again, this is harmonization, which we are not talking about -- but what I don't want to do is ask them to do something that is inconsistent with what CMS is already doing. And so, I would push to move that discussion, that stipulation off, and say that if, and, or when we get the harmonization, the group feels that these things should be harmonized.

CHAIR KAPLAN: Okay.

MEMBER JHA: And leave it at that.

MS. DRYE: The hospital-wide

measure does allow readmissions to be counted as index admissions. Sometimes they are in the same model because it is the same kind, but sometimes they are in different models.

And it is just as Robert articulated. It is just a statistical versus sort of actionability tradeoff. And for this measure, we allow them to be counted. Our other measures we don't. We set a 30-day window, and we don't. But in this measure, for the reasons articulated by the Committee, we felt it was important.

MR. SAUNDERS: We would think of it as a harmonization difference, but --

CHAIR KAPLAN: All right.

Bruce?

MEMBER HALL: Sorry, I'm lost. So, I don't feel comfortable voting on something we are asking them to revise and bring back. So, in my mind, we either vote as is or we say we would like to see revisions and then we postpone voting. And I want a

clarification, what are we doing?

MR. SAUNDERS: And I would think that the planned readmission thing, again, is a harmonization issue. And so --

CHAIR KAPLAN: Yes, I think we need to vote on the -- just because there is such confusion, I am going to ask us to vote on the measure as is. That includes planned admissions.

So, all of the previous discussion -- let's just vote on it the way it is with planned admissions. And if there is call to do that, have those issues resolved in the harmonization discussion.

We are now ready to vote on the measure as proposed by NCQA in terms of its reliability and validity for estimating plans' performance based on hospital readmissions.

MS. ADEELA KHAN: So, scientific acceptability of measure properties was the criterion. Scientific acceptability of measure properties met? Vote 1 for yes, 2 for

no.

And I think the number is 18 we are looking for.

(Whereupon, a vote was taken.)

MR. AMIN: So, Eliot Lazar, scientific acceptability?

(No response.)

Patricia McDermott?

MEMBER McDERMOTT: No.

MR. AMIN: And Mark Williams?

(No response.)

MS. ADEELA KHAN: So, that is 12 for yes and 7 for no.

CHAIR KAPLAN: Okay. So, now we are at almost 11 o'clock. Because it looks like we are going to get into some harmonization questions, for the record, I have to leave at noon. So, you will be absent a Chair as of noon.

So, to broaden the discussion, I really want to be fair to everybody. I would like you to do what we did before, which is

vote on usability and feasibility, recognizing that there are some issues on the consumer side with respect to that, that we really need to be reflected in the report.

So, to parallel the same process that we used for the other measure, I would really like use to vote on this issue first and then come back and ask you to comment for the record, and then put the comments in the report.

MEMBER ALTERAS: I mean, this is nothing different from what I said yesterday, but just to make another advocacy question; that's what I do.

To consumers and purchasers, this is a very usable measure. We don't have a lot of measures that can give us information, and especially since information on what health plans are doing to help give you coordinated care while you are still in the hospital, what they are doing for you after you leave the hospital, how this is going to affect your

quality of life when you go into the hospital.

Again, I know I am repeating myself, but with health insurance exchanges coming onboard, we are hoping and praying that consumers start using this type of information more than ever before, since they are going to have to make all these decisions that they have never had to make before.

And we also are trying to bring into the spotlight what health plans are doing to coordinate care when you are in the hospital, you know, how they have nurses come into the hospital to provide that transition, and to address all the issues of care coordination and care transitions that are driving costs up.

I just think that this is a measure that could really bring light to all those things and find it extremely usable.

CHAIR KAPLAN: Thank you.

Leslie, do you have -- Christine?

MEMBER TRAVIS: I will just say

"ditto" to what Tanya said, but from the purchaser's perspective this is a critical measure that is very usable, and we are already beginning to use this measure at the plan level for accountability and selection.

MEMBER KELLY HALL: I have changed my mind. I guess it is back to usability, too, and it is harmonization, because the confusion never benefits the consumer. It always just increases more confusion and makes any measurement invalid to the consumer.

So, in an effort for harmonization and usability, and then, also, recognizing that markets do vary greatly in the level of involvement of a plan. I am in a State where we legislated out managed care and we are at a 9 percent readmission rate. The plans are not actively involved.

Just an interesting aside, how do we eliminate confusion as we report these things and eliminate burden?

CHAIR KAPLAN: Okay. It is my

understanding that reporting, right, it is reporting an issue for us or no?

MR. AMIN: It is outside of scope.

CHAIR KAPLAN: Reporting is outside of the scope of this Committee. It is clearly an important problem, but it is beyond the scope of this Committee.

So, now I would like us to vote on -- I am going to have to take the Chair's perspective, and I am sorry about going beyond sort of a more full-throated discussion of the usability issue, but I am afraid that we aren't going to get issues, critical issues, involved in harmonization unless we get through this process.

So, let's vote on usability.

Adeela?

MS. ADEELA KHAN: To what extent was the criterion usability met? Vote 1 for high; 2, moderate; 3, low; 4 for insufficient.

(Whereupon, a vote was taken.)

MR. AMIN: Usability vote, Eliot

Lazar?

(No response.)

Patricia McDermott?

MEMBER McDERMOTT: Low.

MR. AMIN: And Mark Williams?

(No response.)

MS. ADEELA KHAN: We're at 16. We need two more people. One more.

CHAIR KAPLAN: Keep pushing.

(Laughter.)

MS. ADEELA KHAN: There we go. We have 5 for high, 4 for moderate, 9 for low, and 1 for insufficient.

CHAIR KAPLAN: Okay. Now we are going to go on to feasibility, make a vote.

And again, Adeela?

As is on the measure. We are voting on its feasibility.

MS. ADEELA KHAN: To what extent was the criterion feasibility met? Vote 1 for high; 2, moderate; 3, low; 4, insufficient.

(Whereupon, a vote was taken.)

MR. AMIN: Feasibility, Patricia McDermott?

MEMBER McDERMOTT: Moderate.

MR. AMIN: Eliot Lazar?

(No response.)

MS. ADEELA KHAN: One more person. Can we just have everyone enter it in again? There we go.

Eight for high, 6 for moderate, 4 for low, and 1 for insufficient.

Does the measure meet all the NQF criteria for endorsement? Vote 1 for yes, 2 for no.

CHAIR KAPLAN: Hang on a second. We are feeling the drumbeat of moving this Committee too fast through this measure, and I want to make sure everybody has time to raise any issues they are finally concerned about, about the overall suitability for endorsement, before we go forward with this vote. So that I am not pushing you guys to run fast on the treadmill and then you bonk.

Jeff.

MEMBER GREENWALD: Just a procedural question. The implications of getting a low usability piece has been the same in the prior as now. What are the implications of that as a Committee? Does that pass, because it's low? I thought if it didn't pass --

MR. AMIN: If it doesn't pass importance or scientific acceptability, the measure does not move forward. The other two criteria should be weighed depending on your evaluation of the overall measure. It could be weighed differently depending on which stakeholder you are, but it is just weighed in the overall.

MEMBER GREENWALD: Because the next question asks, this one says, does it meet all the criteria? And if we say it is low, then does that not pass? Is that sort of definitional?

MR. AMIN: It is not a -- I will

look to Heidi and Helen on this one -- but usability is not a must-pass criteria. So, it can pass overall with a measurement on --

MEMBER GREENWALD: Then I might recommend that we reword the question because, if it doesn't pass all four, it doesn't pass, is how I read that.

MR. AMIN: Yes, that is duly noted. Thank you.

If it doesn't pass here, then it doesn't go on to a continued discussion or, yes, there is no harmonization discussion. It is not recommended for endorsement.

CHAIR KAPLAN: So, it is my understanding, however those go together in your own minds, however you prioritize, and some people may put -- NQF puts the first two as the primary go-forward issues. But if the somehow reshaping of the last two, in your own mind, colors your overall rating, then it does.

Frank?

MEMBER GHINASSI: Procedural question. Given that we are voting on this measure as it stands, and then there is a subsequent harmonization discussion, how does that impact what has already been voted on? Does that mean they are not immutable at that point? Does that meet that, once the Committee has approved it independently, does harmonization imply potential change?

CHAIR KAPLAN: Karen, do you want to --

MS. PACE: Yes. Well, basically, what we say is the final recommendation depends on assessment of any related and competing measures. So, basically, the final recommendations are based on the next step of looking at, are there harmonization issues that need to be fixed in order to make it final or is there an issue of selecting one out of the two before the final recommendation?

MEMBER GHINASSI: So, there is

another vote after this?

MS. PACE: Well, we have to discuss those harmonization issues and what the implications are. There could be a vote; there may not be a vote. It depends on -- so, this will stand as your recommendation unless the issues that come up in looking at the comparison really call into question whether the measures can go forward.

CHAIR KAPLAN: All right. Hearing no other nascent issues before this vote, let's go ahead and vote, Adeela.

MS. ADEELA KHAN: So, again, overall suitability for endorsement. Does the measure meet all the NQF criteria for endorsement? Vote 1 for yes, 2 for no.

(Whereupon, a vote was taken.)

MR. AMIN: Eliot Lazar?

(No response.)

Patricia McDermott?

MEMBER McDERMOTT: No.

MR. AMIN: And Mark Williams?

(No response.)

MS. ADEELA KHAN: We need two more votes. Yes, can you all resubmit one more time, please? The receiver is actually right here. So, if you point over here -- there we go.

So, we have 10 for yes and 9 for no.

CHAIR KAPLAN: We have to go to public comment.

MS. FOSTER: Thank you, Madam Chair, and I will be brief.

Because my comment is to be made generally about the process and the questions you are addressing today, and I am also embarrassed to make it, given the expertise around this table and the seriousness and work that has gone on in this Committee.

But I think that, in fact, this work is being done based on a serious misinterpretation of the Affordable Care Act language. I just want it on the record that,

at least as I read the Affordable Care Act language, CMS must be implementing measures that are condition-specific. It is instructed to look for measures that include all patient data and can be done condition-specific for readmissions, but it is not instructed, nor is it encouraged, nor is there any language in the Affordable Care Act that allows CMS to implement an all-patient, all-condition measure.

And so, I think there is a challenge here to the work and to the interpretation that this work will lead to something CMS can use.

I raise that as a --

MS. PACE: Nancy, could you introduce yourself and who you are representing?

MS. FOSTER: Oh, sure. Sorry.

I am Nancy Foster with the American Hospital Association.

So, I just wanted to lay that out

for the record. And again, I stand in admiration of the work that has gone on so far.

CHAIR KAPLAN: However flawed it is.

(Laughter.)

DR. BURSTIN: And we would be happy to share with the Committee the specific -- and I did find the emails directly back from CMS -- the specific citations they cited within ACA. I want to get them back and forth. But we will share that broadly, including with you, Nancy.

CHAIR KAPLAN: Okay. Now harmonization -- go ahead.

MEMBER HALL: So, Helen, would you mind saying, NQF, do you interpret that way as well? Obviously not.

MEMBER BELLOWS: We actually got information directly from CMS where they cited actually multiple portions of the Affordable Care Act to justify why they needed this

measure and the timeline in which they needed it, specifically wanting to have it in advance of a dry run this spring. So that hospitals would have a chance to review it.

Again, we did rely on CMS's interpretation of the guidance. I did not go and read the legislation in the detail that Nancy did, obviously. But we will go back and clarify that with CMS.

And again, if there is not truly a justified need for expediting it, we will consider whether we can extend the period of time a bit.

But, just to be clear, we were clearly asked to do this in a timeframe with specific citations of three different sections of ACA to justify that timeliness.

CHAIR KAPLAN: I would just like to add, for the purposes of this group, I think that is more NQF's problem. Our considerations are for these measures, what are we looking at with respect to -- whether

NQF is correct or incorrect or somewhere in between about its interpretation or CMS is correct or incorrect, we are just giving our best-possible guidance about the three measures we considered, and the end.

So, now harmonization, but I am informed -- and again, I am sorry, you are going to lose this Chair at noon. But I would like to offer us a very, very, very brief bio and caffeine break to kind of get us to move forward on the harmonization issue, understanding that is really does have to be very, very brief, five minutes. Go.

(Whereupon, the above-entitled matter went off the record at 11:14 a.m. and resumed at 11:23 a.m.)

CHAIR KAPLAN: We are going to start.

And the issues are: travel and logistics, format of the discussion from now forward, travel and logistics handled by Alexis, and then Taroon is going to give us

the agenda for this discussion. And then, we are going to go into a discussion at which time -- and I truly, again, apologize to the group; I do have to leave at noon. So, at noon, I will exit this chair and hopefully we will either have a decision or we will have a replacement Chair by then. Or Bruce Hall will step in.

All right. So, Alexis, do you want to go?

MS. FORMAN MORGAN: Yes. So, just very quickly, for those who want to rebook their flight for an earlier flight, if you can call our travel agency -- and here is the number posted -- they will work with you to rebook your flight.

MR. AMIN: Okay. So, the format of the discussion going forward is that what we would like to do is identify the specific differences between the two measures that were recommended to ensure that they are harmonized.

There are four specific issues that were identified during the past discussion, but we will obviously welcome the Committee to identify further differences that should be clarified.

What we would also like do during this discussion to facilitate and reduce the number of conference calls that we would have to have subsequent to here is to really identify which one is preferred and get an agreement in the Committee on which approach is preferred.

To help facilitate this discussion, in the prep materials that Alexis had sent out there is a side-by-side table of competing measures that we could -- I am thinking of the best way. Okay. So, we could also post that. We will figure out the best way to run that.

So, the four issues that were identified during the previous discussion were the handling of planned readmissions, whether

or not a readmission can be identified as an index admission for subsequent readmissions, the handling of behavioral health in both of the measures, and also an issue that was raised in the risk-adjustment group methodology. And the condition category, one uses HCCs; another uses CCs.

So, if we just clarify the approach in each and see if we want to understand why there is a difference and whether that difference should be justified in moving forward.

And I will turn it over to the Chair to identify other major harmonization issues and then facilitate a discussion on preferences of the Committee.

The planned readmissions, subsequent readmissions serving as an index, HCCs or CCs, and behavioral health.

CHAIR KAPLAN: It is my understanding from the group -- and correct me if I am wrong -- the measures developers do

not have to agree on the details of how they are work this out. They just have to agree here that the will work together to solve this problem. Is that accurate?

MR. AMIN: Well, we would like, if there is a difference in the methodology, what we would like to do is identify which one we are going to forward with. Ideally, they would be harmonized across the two measures. That level of detail would be needed, so we can cut down on future work.

CHAIR KAPLAN: Okay. All right. So, we are putting these in the blender. Both are going into the Cuisinart and we have to come out with some merged product, which at least, although the exact details of how they do this don't necessarily have to be worked out with us, the issue has to be somehow resolved that they can. Is that --

MS. PACE: Right. I think one of the things that hopefully will be facility this is that both of these measures are under

contract with CMS. So, to a certain extent, it would be in everyone's interest to have harmonization to the extent possible.

So, what we would like to do is identify the issues. If there is a preferred approach, obviously, they are going to have to get together and with CMS decide what is practical across both measures. And ultimately, then, if there are any issues that can't be resolved at this stage, what the plan would be for the future.

So, I think to a certain extent they are going to have to respond back to us, not today, but we want to at least identify the issues, and if there is some kind of preference, and knowing that they are going to have to look at that and see whether that can actually happen.

CHAIR KAPLAN: So, I just exercise the Chair's priority, and I hope doing these in reverse order from the easiest to probably the most controversial.

Oh, yes, okay. So, now we need to add anything that is not here that is burning a hole in your pocket.

Brent? Bruce? Bruce and then Brent.

MEMBER HALL: Just issues that need to be addressed. I think that structured cohorts and straight logistic versus a hierarchical approach.

CHAIR KAPLAN: Brent?

MEMBER ASPLIN: I would make an argument that they are competing and that we shouldn't even be talking about harmonization. I don't feel incredibly strongly about it. We can certainly just move forward with the discussion. But, really, it comes down to, what are your thoughts about accountability; what are your thoughts about the populations that the measures address, and how do we want the public to respond to that?

I was really struck by Tanya and Sherrie's comments about -- or excuse me --

Christie's comments about the enthusiasm for the usability of the NCQA versus the hospital-based one. I actually think in some cases shared accountability is no accountability. And I think if we need to make progress on readmissions, it is delivery system responsibility.

If you have both of these measures active in your marketplace, I think there is going to be a lot of confusion because they don't address the same populations. There's no fee-for-service Medicare in the NCQA one. There's no Medicaid managed care in the NCQA, which could be. And I would argue that they are competing.

Sorry to throw a curve ball, but that is where I would sit.

CHAIR KAPLAN: Yes, we had a similar discussion last night.

Jeff?

MEMBER GREENWALD: Just a quick question for both developers maybe. Is the

handling of cancer patients the same in both models? It was clearer in one than the other to me.

MR. SAUNDERS: It is not. We include the cancer patients, and so their planned readmissions would be included. And so, it is something we can address by removing that population. There are a variety of ways to implement removing that, but they have made that exclusion.

MS. DRYE: Right. Specifically, we exclude from the cohort patients admitted for medical treatment of their cancer. Patients who have cancer admitted for surgery, for example, are in the measure.

CHAIR KAPLAN: Anyone else?

MEMBER KELLY HALL: I would go back to Bruce's original question and ask the developer or Elizabeth to comment. When you asked, "Can this be used," she said, "Not today." Then, what would be required to have a single measure? What is the gap that you

feel that exists?

MS. DRYE: The gap is, well, we have tested the measure and provided to NQF the measure's -- the measure works well in the 18-and-over population. So, we just provided that in the last week.

So, the gap is actually in just having data to test it. We don't have health plan data to test the measure, and I think it would be important just to demonstrate it. But our expectation is that it would work.

MEMBER KELLY HALL: Thank you.

CHAIR KAPLAN: Tanya?

MEMBER ALTERAS: Yesterday, when I left here, I thought that these measures really couldn't be harmonized, but I didn't really see them as competing. I saw them as two different measures in two different settings. I could see them both co-existing in implementation because I didn't actually see them being reported in the same place or being used by the same people.

But, after today's discussion, I do feel like they could be harmonized, especially after hearing the responses from NCQA, and the fact that they are both funded by the CMS definitely helps.

And what Richard was saying about the burden on hospitals to provide data in different ways, well, it seems like there are ways that that could be streamlined and aligned.

So, I am definitely in the camp that sees this as a harmonization issue and not a competing measure issue.

CHAIR KAPLAN: Let me just ask you a question because I think, unless I have gone completely daffy in this discussion, that you did raise the issue of confusing consumers, and that having both was a problem.

Or you raised the issue of confusing consumers? Okay.

So, you don't think it confuses consumers?

MEMBER ALTERAS: Well, I see the Yale/CMS measure being reported on Hospital Compare. I don't see the NCQA measure being reported that way, unless they were harmonized in some way, because Hospital Compare doesn't report at the health plan level.

MR. SAUNDERS: We would use Plan Compare.

MEMBER ALTERAS: Yes. So, I don't see the confusion there because I just think that they would be used in different venues.

MEMBER KELLY HALL: As health information exchanges and health data exchanges, or HIEs, are being merged, I think we will see single sites being used by consumers because the economies of scale in the states, as they look at trying to fund both of those data consolidation efforts, you are seeing them come together more.

So, I would think that it is a matter of time where you will see comparison as a payer and a provider being considered

either in a plan selection or a provider selection both.

MEMBER ALTERAS: Yes, and let me clarify then. And I agree. I mean, I have been the one tooting the health insurance exchange this whole time.

But I just think it is different levels of decisionmaking. I think you would use the NCQA measure when you are trying to decide your health plan and use the hospital-level measure when you are going in for something else, when you are actually choosing your hospital. I know on the hospital side the accountability is different as well. But I guess I don't see this as being a confusion issue.

CHAIR KAPLAN: Okay. I am going to pull us back to, does anybody have anything that is not on the list right now that needs to be on the list? Because I really want to make sure that the topics get covered rather than getting into the discussion ahead of

time.

Laurent?

MEMBER GLANCE: I am going to ask for the group's forgiveness because I think the point that I am going to make may not be adding one particular category.

But when I look at this, the big-picture to me is that we are trying to fit a round object in a square hole. I think that when you have an outcome that you want to risk-adjust, you want to have one single risk-adjustment model. And you can use that model at different levels, whether it be at the hospital level or whether it be at the plan level.

The idea that we should be trying to take two really very different models and somehow getting these folks to sort of come together and make them look alike, I mean we should just have one model. Unfortunately, the forum that we have used today to sort of discuss these models hasn't really allowed us

to sort of decide which model is better. And we may not even have enough data to decide which -- I don't think we do -- to decide which model is better.

But these are two very different models. You have one model which is a single model, another model which incorporates five condition-specific models. You have one model that is hierarchical; the other one is non-hierarchical. And they have different respecters.

And we know, and there is a lot of empirical data out there, if you use different models, you will come up with different results in terms of which hospitals are high-quality and which hospitals are low-quality.

And I think -- and I am going to finish right now -- the point that Richard made earlier was a really important one. You are a hospital; you have one primary plan, and you get two different grades. That is going to be very confusing and is really going to

kind of, I think, affect the face validity of what we are doing here, the whole benchmarking enterprise.

And I don't have a solution, unfortunately. But I think it would be best if we had one measure, one model, not two different competing models.

CHAIR KAPLAN: Remember that, again, just to make sure that we stick to the sort of measure itself, rather than how it gets reported. I know that your issue is a little broader than that, but I wanted to rein it in here, to not have reporting issues come up.

Okay, Frank, is this something that is not on here yet?

MEMBER GHINASSI: It is on here.

CHAIR KAPLAN: Okay. Jo Ann?

MEMBER BROOKS: It just kind of adds onto this comment. If we ask, let's say, NCQA to make changes into the population that they are planned, to take those planned

admissions out, isn't that, then, changing the psychometrics of their model that it was built upon? So, we would be basing it on something different than what we have that we voted on.

CHAIR KAPLAN: Robert?

MR. SAUNDERS: I agree that if we change the specification that it is different. But I think that is the nature of this process. You voted at one stage. We are at the harmonization phase. We will try to get agreement. And then, you will vote again after that: have we harmonized enough, and that you go forward with both measures.

So, is your question, then --

MEMBER BROOKS: I guess my question was just, we looked at the data, the psychometrics of your model as it was initially developed. And now, if we change some of those parameters and the population, taking out the planned, et cetera, that were in your model development, does that change the psychometrics and make it a different

model?

MR. SAUNDERS: It clearly makes it a different model, and I would think that you would have to -- I would say you would have to. I mean, I think you would want to look at the results of that model, and I think that would be a fair issue to sort of say, when do you need to see that in order to make your decision, and whether we can do that. We can certainly modify the specification.

MEMBER BROOKS: Yes, and I guess my question was not to hold things up, but just that, in fact, that would change the model, and we would need to look at that, and that has to be considered.

CHAIR KAPLAN: Okay. Let me intervene here a little bit because I think we have an issue to resolve that has a procedural quality to it.

So, do you want to, Taroon, take over and tell them what they need to do?

MR. AMIN: Okay. I know the

Committee is probably frustrated, thinking that they have made final recommendations a number of times already.

(Laughter.)

However, the final recommendation is based post-the-discussion around competing or harmonized measures. Unfortunately, actually, Brent isn't here for this part of the discussion because he is actually the one that raised this.

But the Committee will first have to evaluate whether or not these measures are competing. And if they are competing, there may be a justification for why you do need two measures, a conversation that Christie brought up around really needing it potentially for purchasers, or arguments of that nature.

There may be reasons why you need two measures. In that case, measure components would need to be harmonized, which is sort of the kind of discussion we are getting at here.

We are really, first, as sort of the conversation will go, we will first have a discussion around whether or not they are competing, have a vote on that. If they are competing, if they are defined as competing, as having the same measure focus and have the same target populations, then there will have to be a selection of best in class, where you will select one of the measures as recommended to go forward.

If they are competing, and you still feel like there needs to be two measures because they are measuring two different levels of analysis, and there's a justification based on the Committee's deliberations, then we will decide what are the components that need to be harmonized across the two different measures.

Helen, Heidi, others that want to sort of procedurally clarify this for the group?

So, Brent, this is a direct

response to the question that you are asking. So, there will be a vote on whether or not they are truly competing measures. And based on that, there may be a justification, as I just explained.

CHAIR KAPLAN: So, how many are confused about what we are voting on? A lot of confusion.

Okay. So, here's the way I understand this: there's three possible options. One is there is a best in class, which means there is one class. So, there is one class of measure and there are two possible measures as awardees for that best-in-class designation.

Another consideration is there are two classes, in which case we don't need to vote on anything because we have two measures, one in each class. Correct?

MR. AMIN: A best-in-class definition would mean that these two measures would compete head-to-head and you select one.

I just want to make sure; procedurally, yes, that is correct.

So, if you decide that they are competing, you will select one for final endorsement. If they believe that they are competing but you feel that both measures are needed, then there would need to be a justification and a discussion around how these two measures will be harmonized.

And if they can't be harmonized at all, then they just stand as they are.

CHAIR KAPLAN: Yes. So, I think you just heard the three options, I think, and let me try to get them right.

MR. AMIN: Okay. No, please do.

CHAIR KAPLAN: Okay.

MR. AMIN: I want to make sure everyone is clear on this.

CHAIR KAPLAN: So, best in class is there is one winner, right? So, there is a single class; there is one winner, period; the end.

The second consideration is there are two possible classes because these aren't competing; therefore, they represent two different distinct, but related issues, in which case --

DR. BURSTIN: So, basically, they are competing is a yes/no. And then, if you decide that they are not competing, they still are likely related. It is the same measure focus, essentially, potentially the same or vary in population. And then, you would get into the issue of harmonization.

So, your first discussion, I mean, really, to Brent's point earlier, is: are they competing? Should there be only one of these? And if so, which one? And if not, we will get into the harmonization discussion.

MR. SAUNDERS: I would draw the decision tree if we had a piece of paper.

CHAIR KAPLAN: So, let's just do this in phases, so I don't get messed up.

Are these competing measures, is

our first vote. We are going to do it in the Adeela specialty way of posting the vote, and we are going to vote on this issue.

And Bruce has an issue.

MEMBER HALL: So, I don't feel that these measures are competing. In terms of reporting at different levels, I don't feel they are competing. Where I feel they are competing is they are putting forward two different methodologies that creates an undue and unjustified burden.

I don't have a problem to have a measure that reports at the plan level and a separate measure that reports at the hospital level. But we have highlighted six or seven differences in methodology. That is where I feel they are competing.

And I would like to say, tell me why they have to be different on that method? I don't mind that you are going to report one on Plan Compare, report one on Hospital, and so forth. But this creates a burden for

consumers to understand and for hospitals to pull off or plans to pull off.

So, I feel they are competing methodologically, and that is what I would like the developers to comment on. In a prior NQF project where there was a similar issue, the decision was have the developers either harmonize the approach and say, "You're right, we didn't need two different approaches. We've decided collaboratively on one approach to that," or justify the necessary difference between the approaches. That is my perspective.

MR. AMIN: So, in that, the logic of how that will play out is, in that case, you would vote that the measure is not competing, but they are related and the components of the measures should be harmonized. And all these components that we were discussing, we will have a discussion around which should be harmonized and actually potentially selecting one versus the other, if

there are two.

So, that is how that scenario would be voted through our process. Is that fair?

CHAIR KAPLAN: So, you put them in the blender if they are related, and there's two separate blenders if they are not and they don't need to be blended because they are two really distinct things.

So, the first vote is, are these measures competing, in which case we put them in the blender and make them harmonized, or are they not competing and they are measuring two really different things?

No. I just created a -- okay, so go back to whatever it was that was the issue beforehand, and vote that way.

So, reiterate it Taroon.

MR. AMIN: Okay. Competing means you have to select the best in class and there is only one at the end of this project.

If they are not competing, they

are related and need to be harmonized. And the components of harmonization are before you, and they will be entertaining -- you know, we will all discuss this. And then, we will select which components need to be harmonized, in a very short period of time.

(Laughter.)

Okay.

CHAIR KAPLAN: Okay. Apologies for the -- my excuse is I have no excuse.

So, are these measures competing, yes or no, is how we are voting. Is everybody ready to do that? Are they competing, yes or no?

So, Adeela?

MS. ADEELA KHAN: So, are these measures competing. Vote 1 for yes, 2 for no.

(Whereupon, a vote was taken.)

MR. AMIN: Eliot Lazar?

(No response.)

Patricia McDermott?

MEMBER McDERMOTT: Yes.

MS. ADEELA KHAN: So, we have 7 for yes and 12 for no.

DR. BURSTIN: I actually would like to have Bruce restate what the basic question was because that actually might simplify our issues a bit. The two questions that I think you asked might be a way for us to kind of work this through.

MEMBER HALL: My most recent comment was just that I didn't think that they are competing in terms of reporting or reporting level. I feel what they do is they create an undue burden to calculate something two different ways.

If there is a very good reason, an absolutely necessary reason for any of those six or seven items, then that is what the developers have to tell us. But if there's not, then they should bring back an approach they have agreed to do collaboratively and in a similar methodologic fashion.

MEMBER KELLY HALL: I have a

question. So, do we, then, have to state what bias we have of the methodology and the biases we have, so that we don't come up with something that is blended in the blender, but now a third option or now something that was against the bias of the Committee?

For instance, if we say that planned readmissions need to be excluded, we have said that, and that the group comes back to something now new, how do we incorporate a bias in this process?

CHAIR KAPLAN: I would think it is our discretion to say we would like comments from the developers on these issues, and we have a bias that planned should or shouldn't be included on any one item. But, beyond that, it is up to the developer to come back with a response.

DR. BURSTIN: But I think the Committee, then, needs to state a preference on some of these before they go back to the blender.

MR. AMIN: So, I am attempting to put all these considerations into a few buckets, and then we can sort of walk through them individually.

We have three major risk-adjustment issues: the HCCs or CCs, the logistic versus hierarchical, and the structured cohorts.

For the inclusion criteria: the handling of planned readmissions, using readmission as a subsequent index, hospitalization for subsequent readmissions, the handling of behavioral health patients, and then there's exclusion of cancer patients. And this one is repetitive. So, I will take that one out.

Are there any other major concerns that the Committee wants to raise? We will make sure they get on there, and then we could start to state preferences and have some discussion around that.

So, I will turn it back to

Sherrie.

CHAIR KAPLAN: So, now I am even confused. We are talking about the possibility of now -- we have decided these are not competing measures, correct? Okay. So, now they need to be harmonized.

All right. So, if these are not competing measures, then they are measuring the same animal, and now they need to be put in the blender and we need to resolve these issues.

So, I mean, I think the best way to do this, unless somebody has a different preference, is to start at the top and work these through one-by-one, unless you have got an issue that you think is probably easily resolvable.

So, go ahead.

MEMBER BANKOWITZ: I think this last one, exclusion of cancer patients, none of the models excludes cancer patients. One tries to exclude cancer patients that are

admitted principally for cancer, which is related to planned admissions. So, to me, it is a subset of the plan. Maybe that is a nuance we don't need to deal with, but --

MR. SAUNDERS: Yes, that was my interpretation.

CHAIR KAPLAN: Good point.

We want to start the discussion on risk adjustment and the assessment of comorbidity.

Do you want to handle this so that we ask the person who raised the question to -- all right, I am just asking the measures' developers to tell us whether this is a resolvable issue?

DR. BURSTIN: I'm sorry, I think at this point we just want to get a sense of the Committee: on any of these issues, is there a clear stated preference one way or another? So we can give that information to the developers as they do their work.

MS. PACE: And understand that

they are going to have to really look at these and see what is possible. You know, we have heard a lot about these individual issues in the context of these measures and this timeframe. So, again, it is going to have to go back to the developers to really then examine what is possible and preferable, given the two measures.

MEMBER ALTERAS: Okay, I will start. I would just say my preference is against hierarchical risk modeling, if there is going to be harmonization on that bullet. The risk adjustment, the hierarchical risk adjustment versus logistic. I am looking at the second bullet.

DR. BURSTIN: That is one of the differences.

CHAIR KAPLAN: Okay. We are going to shift leadership. Thank you all, again, for all your hard work. It really was a privilege to be able to sit in on this discussion. And with so much talent in the

room to address this issue, I think there stands a reasonable shot at coming out with something very, very good.

(Laughter.)

Thank you again.

ACTING CHAIR HALL: So, I will just suggest a default that we can work off of. The default would be that this Committee would be agnostic to stating biases, but that for each item we want the developers to convince us of their ultimate decision.

But there was just one comment where there was a clear preference for a bias in one direction, and so that is what we are looking for right now, is any clear preferences that would veer from the default, the default being the developers convince us of their ultimate decision.

Anything you want to bring up from the whole list? Ashish?

MEMBER JHA: So, I am going to say a couple of things that I think reflect the

broad view of the Committee. But I think planned readmissions, there should be some effort to get rid of them from the NCQA effort. Do they need to be perfect in terms of exactly the same patients excluded? Ideally, yes. If we can get very close, that will be a big step forward.

I am also going to put in that, on the index readmissions, I actually misunderstood what CMS is doing, I mean what the Yale group is doing. But the preference would be what Jeff suggested, which is that readmissions should be able to count as an index hospitalization for further readmissions.

And I guess maybe one last thing, which is I know there was a preference for logistic. You guys probably guess where my preference on this is.

I would only argue that we don't end up with hierarchical models for both. I think it is unnecessary at the plan level.

So, if the decision by the developers is to stick with hierarchical model for the hospital, that is one place where I can live with a divergence between the two.

ACTING CHAIR HALL: And again, that would be something where the developer would come back and say, "Here's a difference that we think deserves to be preserved for the following reasons."

Leslie? I'm sorry. Frank?

MEMBER GHINASSI: Thank you.

I would also support the concept of including the structured cohorts in a homogenized model.

And second, as I had indicated yesterday, the reason stated for excluding behavioral health was that most of these folks are treated in specialized psychiatric facilities or rehab hospitals. I don't agree with that statement and do not believe that they should be excluded. Many, if not the majority of patients, are treated in acute

care facilities.

I just don't see a reason for excluding that. I think it would be a disservice to that whole population group, and I also think it is a key comorbidity factor. It ought to be able to be studied both as a primary diagnosis and a comorbid diagnosis.

ACTING CHAIR HALL: And so, if the developers were to come back with preserving the exclusion, we would want to see that further justified, the exclusion justified. Frank is listing our default as inclusion. We would like to see better or further support for excluding.

Richard?

MEMBER BANKOWITZ: I support the recommendations. The last one, I think there could be a structured cohort in itself.

The one recommendation I heard, that we might have one logistic and one hierarchical, I am uncomfortable with. I think that the hierarchical is trying to

account for hospital effect, and I would like to move away from having two different results. The closer we can get to one unified method, the better, and we should try not to diverge on this.

ACTING CHAIR HALL: Thank you.

So, Taroon, if you could, under behavioral health, just add there "consider cohort". And then under HCC versus CC, would you put "default, one approach", "single approach"?

Jim?

MEMBER BELLOWS: Yes, thank you.

Did we lose the cancer thing further down in there, medical treatment of cancer? Well, I want to get it back on the list because it remains a difference. But I think from a perspective of many users, exclusion of medical treatment of cancer is problematic, and I would like to see it re-included in the direction of the NCQA measure.

MEMBER ASPLIN: Jeff?

MEMBER GREENWALD: I am just wondering as a process if this is a potentially not-very-efficient approach. I am worried that, if we don't go through these sort of line-by-line and get a group consensus around this, we are going to give them very mixed messages.

And so, I wonder whether it would be better to just sort of go line-by-line and say, "You guys figure it out" or "We think it should be this way."

ACTING CHAIR HALL: Okay, well-taken. I think that is a great point. So, now that we have got some thoughts up there, let's do exactly that.

Under risk adjustment, the two separate approaches to structured data, does anyone want to add or change to the notion that our default would be single approach, unless the differences are well-justified? Does anyone want to add to, modify that?

Is this what you had in mind,

Jeff?

MEMBER GREENWALD: Yes, and I don't know if we can do the voting or not, but I am just worried that there will be a few people speak up and we have potential conflicting, and that is going to be very mixed messages to our users.

So, I would suggest that we propose sort of an A, B, or C. Either we keep A, we keep B, or we throw it back to them to figure out how to make it work together or justify not working together, as a third option.

ACTING CHAIR HALL: Would everyone be comfortable with us saying, you know, our real default is that we are agnostic and we are trying to indicate a potential bias, and that we are still going to rely on the developers to say, "We think that bias you put up there is just off the wall."

Is that close enough, Jeff, or do you think for each line you want to vote on

three options?

MEMBER GREENWALD: I'm flexible. I think the three options is just the least cognitively-challenging for my simple mind.

ACTING CHAIR HALL: Because I don't know that we are going to be doing this sort of automated response voting. We can do show of hands or we can just do our best to say, you know, be courageous; speak up.

Under HCC versus CC, we are sort of saying our default is a single approach, but speak up if you have issues with that.

MEMBER KELLY HALL: I have a question.

ACTING CHAIR HALL: Yes, Leslie.

MEMBER KELLY HALL: And I don't mean to be inflammatory, but we have one group -- maybe I am mistaken -- that has intellectual property associated with a measure and one group that does not. Does that matter?

Okay. Thank you.

ACTING CHAIR HALL: So, either I need procedural guidance on whether people are comfortable or want to try to respond more to Jeff's concern, or is there a separate point of input?

I am very sensitive to what Jeff said. Let's see if we can go line-by-line, indicate any minor or major bias we might have. But, at the end of the day, we are relying on the developers to make comments on each of these points.

Jeff, please tell me if I am not addressing that adequately.

So, anything else on the first line, HCC versus CC, and the attendant issue?

Richard?

MEMBER BANKOWITZ: I would add a recommendation that the developers try to account for hospital-acquired conditions.

ACTING CHAIR HALL: So, provide us with information on prior-to-admission handling?

MEMBER BANKOWITZ: Handling those things that occurred in the hospital as a complication of care.

ACTING CHAIR HALL: So, some additional comments on that perhaps.

Laurent?

MEMBER GLANCE: I would just like to point out that what we are doing here is we are calling this harmonization, but what we are really doing is we are asking the measure developers to go back and really completely revamp their models. This is very different from what we talked about the first day, where we said: look, you've got to kind of take the models as they are. I mean you sort of vote them up and down. But your job as a Committee is not to send these folks home and redo the entire models.

I think that what we are doing right now is just that. We are changing the very nature of the models.

Before we go on with this process,

if this is what we want to do, I would just like to hear some comments from the measure developers themselves on how feasible this is going to be for them to go back and go through this process.

ACTING CHAIR HALL: So, I think what we are doing is we are putting some topics up for the measure developers to provide feedback to us. It will be within their authority to say, "That's not something we are willing to or able to change."

On your first point, we have voted on them as they are and we have reached a decision as they are. But I do believe -- and the NQF staff and leaders can correct me -- I do believe it is within the realm of authority for the NQF to say, "You've been approved, but we are demanding some attention to the following issues."

I have been in a project before where this has happened. So, I do believe that is within authority or within scope.

DR. BURSTIN: And that is certainly a change from a couple of years ago, when you were last engaged with NQF. So, we actually now are actively pushing developers after measures are both approved to actually change their measures and harmonize.

Actually, Bruce is referring obliquely to a rather painful process of the CDC and the American College of Surgeons to actually combine their surgical site infection measures into one. So, this is some precedent. We have been really trying to say let's not create cacophony out there. Let's try to actually be value-added.

ACTING CHAIR HALL: And even on that background, again, the developers have the right to come back and say, "That's not feasible. We are not willing to change on that." And then, this Committee would have to reconsider what they said.

Do we want to move down the list? Logistic versus hierarchical, I think we have

-- I'm sorry -- HCC versus CC, are we set?

MS. DRYE: Do people want to know, just a clarification, what that difference is?

ACTING CHAIR HALL: Sure.

MS. DRYE: Okay. We both use CMS-maintained grouper for the ICD-9 codes that groups 15,000-plus ICD-9 codes into condition categories. There is a hierarchical component of that that is used in payment. And NCQA uses that hierarchy. After you have grouped conditions into clinically-coherent categories, they apply a hierarchy.

We don't use it because we are really interested in just grouping those codes into clinically-coherent groups. When you apply the hierarchy, things happen like your hypertension gets cancelled by a higher essentially cost condition. We don't want those changes happening, and we want to be able to report the frequency of the risk variables in each hospital's population to the hospital, and the application of the hierarchy

changes those.

So, we use the condition categories, but we don't apply the hierarchy. NCQA applies the hierarchy. That is just in essentially taking the ICD-9 codes and accumulating risk-adjustment variables.

ACTING CHAIR HALL: Thank you. That is very helpful.

So, I think that would be the sort of issue around which you could make an attempt to either convince us to preserve your differences or agree to treat it one way or the other.

On logistic versus hierarchical, it is listed as preference logistic. Is that how the Committee feels? Or preference, single approach? Or how would people prefer to list that?

Laurent, are you still up? Leslie? Richard?

MEMBER BANKOWITZ: I would prefer a single approach. I don't think we should

dictate which one it should be.

ACTING CHAIR HALL: Single approach?

Jim?

MEMBER BELLOWS: I guess my preference would be preserve logistic, which means either logistic or both, but not a single approach that is hierarchical.

ACTING CHAIR HALL: Okay. So, I think we have -- Ashish?

MEMBER JHA: I was going to basically say what Jim said, which is I don't think there is any reason to do hierarchical for the NCQA measure. So, either we split it or we go logistic for everybody.

MEMBER TRAVIS: And I would agree with that.

ACTING CHAIR HALL: Okay. Is everyone comfortable listing that as a bias that they can respond to? We are dictating the response necessarily yet.

Okay. Then, moving on to

structured cohorts, the notion that there should be some division, separation of models in the analysis versus not. It looks like there is preference for inclusion that is listed so far. I don't know if that reflects people's will.

Any comments?

(No response.)

Okay. We will let them respond to that.

Inclusion criteria: planned. Have we hit that already? Any additional modifications or comments?

(No response.)

I'm not seeing anything.

And including the cancer issue?

MEMBER JHA: Can I just get a clarification --

ACTING CHAIR HALL: Yes, Ashish.

MEMBER JHA: -- because I think it is a little complicated. So, I think where we came out as a Committee, but please tell me if

I got this wrong, is, in general, we think planned readmissions should be excluded. However, I think what we heard earlier was that the cancer patients should be included. So, planned readmissions excluded; however, cancer patients and that cohort included. That is the default preference of the Committee.

MEMBER BANKOWITZ: It is not my preference.

MEMBER JHA: It is not? Okay.

MEMBER BANKOWITZ: No. Because I think we should understand what that decision does to the model. If it includes quite a few planned admissions, then I don't think we should insist upon it.

ACTING CHAIR HALL: Yes. So, I mean, our option is to kind of remove what we are stating any bias might be and let them respond.

Mark?

MEMBER SCHUSTER: I thought there

was a way to do both. I mean, cancer patients don't be kept out completely, but chemotherapy, repeat visits, are left out. So, I thought --

MS. DRYE: Yes. Can I just clarify the difference? Because this happens in our own discussions all the time.

When we are using inclusion, we are really talking about what index admissions are we evaluating for readmission. We don't include index admissions for cancer patients admitted for medical treatment of their cancer. There are many cancer patients in the measure otherwise.

The issue of planned readmissions applies to how we are defining the outcome. And so, there we specifically have an algorithm, and we don't count admissions for chemotherapy. I mean we call those readmission planned. So, they do not count as a positive outcome in the measure.

ACTING CHAIR HALL: Right, right.

Thank you. Thank you for that clarification.

And I would just reiterate, on behalf of the other measure developer, something that they stated before, which is, in their approach, using the categories that they use, they will capture some endogeneity of readmission to a particular diagnosis. And so, some of that is controlled for.

But, again, where we are is, are we comfortable with what is up there for them to respond to or not?

Jeff?

MEMBER GREENWALD: I guess this is a question for the Committee. Are we asking them to -- I think it is fairly clear the message is we need to work on excluding planned readmissions.

One group has a methodology for that. Are we asking them both to adopt the same methodology or would it be acceptable if the NCQA group came back with a different methodology for excluding planned

readmissions?

ACTING CHAIR HALL: I think we are asking them to respond to this concern. I don't think we are dictating what the response has to be.

MEMBER GREENWALD: I guess part of my concern would be, if they came back with a different methodology, we would then add to the confusion question again.

ACTING CHAIR HALL: No, falling back on the notion that the two measures have been approved as is, we are not asking them to independently come back with two measures that are not harmonized. We are asking them to either say this is our new approach to harmonization or we are justifying our difference.

So, Mark, are you still up?

Okay. On the planned readmissions, is there anything more we want to add?

(No response.)

Not seeing anything, the index readmission, we have sort of discussed. Anything to add? Handling what is called an eligible index.

(No response.)

Behavioral health, any further comments?

(No response.)

Does the list go on? Or is that the whole list? Okay.

Jeff, more comment?

MEMBER GREENWALD: Yes, just on the behavior health side of this, I also want to make sure that how we are identifying substance abuse admissions is similar between the two groups and that they are both being handled in the same way.

ACTING CHAIR HALL: So, perhaps they could comment on that or explain to us any difference.

Okay. So, any larger, bigger-picture comments?

Again, what we are saying now is both measures have been approved as is. We feel very strongly as a Committee that important attention should be turned to harmonizing the approach to measurement and calculation, including, not necessarily limited to, these issues.

And we would like the developers to respond collaboratively to this list. Respond collaboratively means respond to us. We are not dictating whether that means one changes, the other changes, neither changes, or what.

They come back and they say, "Here's something we have agreed to do the same way" or "Here is how we are justifying this difference, and we think it is important to preserve the difference."

Any additional comments? Yes, Tanya?

MEMBER ALTERAS: Just a quick process question. So, before this goes out

for public and member comment, we are going to get to see the responses, right? Okay.

ACTING CHAIR HALL: Jeff?

MEMBER GREENWALD: Sorry. I want to just spend a second on the substance abuse question, since we haven't gotten over this yet.

It is a major predictor of readmission in many populations studied. And so, the question is, does this group recommend that it be excluded, as some of the protocols have?

ACTING CHAIR HALL: When you say "this group", you mean our Committee?

MEMBER GREENWALD: Our Committee.

ACTING CHAIR HALL: Okay.

MEMBER GREENWALD: Because it is a different kettle of fish for sure in terms of the interventions that are appropriate. But from a hospital and health plan perspective, it is a major driver of cost and utilization.

And so, the question is whether

excluding it would then take out of the limelight a major utilizer of resources and a potential opportunity to improve quality of care.

ACTING CHAIR HALL: Other Committee members want to comment?

MEMBER JHA: I would keep it in. I don't know --

ACTING CHAIR HALL: Jeff, you are in favor of keeping it in as well?

MEMBER GREENWALD: I am in favor of keeping it in.

ACTING CHAIR HALL: Okay.

MEMBER GREENWALD: And I recognize that that is potentially problematic at the end-user level, but it also, I think, has some serious benefits that I think from the national perspective we ought to be acknowledging.

ACTING CHAIR HALL: And, Ashish, you support that?

MEMBER JHA: I do. And I think it

is going to be less of an issue for the Yale model for now because it is people over 65, but when people decide to apply that for a broader population, it will be more of an issue. But I do think, in my mind, we should keep it in.

ACTING CHAIR HALL: Okay. So, at the moment, we are just asking for a response on this topic.

Leslie?

MR. AMIN: Bruce, can I just clarify?

ACTING CHAIR HALL: Yes.

MR. AMIN: The CMS measure, additional testing has been provided to extend the measure to 18 and over. It is not an over 65. Okay. I just wanted to make sure.

MEMBER KELLY HALL: I have a naive question, I'm sure, for those expert in the psych area.

But so much of psychiatric care is not covered by a plan. Do we have any issues

with the fact that that plan data doesn't have it and it is now comparing against hospital data that might? I mean, is that part of a harmonization issue? Or the plans don't have to report something they don't cover.

ACTING CHAIR HALL: I am not an expert on this. So, I would say at the moment we are asking the two developers to comment on this issue.

I think Richard was next.

MEMBER BANKOWITZ: I do recommend, if we keep it in, that it become a separate cohort because the comorbidities in that group, the risk adjustment for that group is quite different. Comorbidities don't apply to that group in the same way. And so, I think that if we are going to compare, we need to compare apples to apples.

ACTING CHAIR HALL: Jim?

MEMBER BELLOWS: I'm sorry if I am the only one who lost the thread, but I am not clear on whether we are only talking about

inclusion of admissions and readmissions where substance abuse is the principal diagnosis or whether we are talking about substance abuse as a risk adjuster for all readmissions. And if so, I am worried about gaming and data capture and completeness and all kinds of other things that we haven't begun to talk about yet.

ACTING CHAIR HALL: I think we were talking about whether to include these admissions as eligible index admissions. Is that correct?

All right. Leslie? Okay.

Would each developer mind in one sentence restating how they handle behavioral health admissions?

MS. HORWITZ: So, I want to clarify how we handle this because it is actually a little more complicated.

In our dataset we have 21,483 of the 8 million admissions that are for purely psychiatric admissions. And by that, I mean

schizophrenia, mood disorders, personality disorders, adjustment disorders, the kinds of things you would be admitted to a psych hospital for.

And the reason we have so few of those is because we have sort of an acute care hospital measure and we just don't have data for those. However, we have many, many, many, many, many patients with substance abuse disorders with alcohol dependence, with alcohol withdrawal, with all of those things, in our measure, because they get admitted to acute care hospitals, as you said. And we have them; we put them in our medicine cohort because they are typically cared for on medicine units.

So, we do include all of those patients. The small subset that we excluded are the patients really admitted for true psychiatric treatment who are generally admitted to psychiatric hospitals, and a few, 21,000, somehow slipped into our 8 million

admissions.

And that is why we sort of report that we have a psychiatric exclusion. But I think the majority of the patients that you are talking about actually are in the measure.

ACTING CHAIR HALL: Robert?

MR. SAUNDERS: We include psychiatric disorders as an index condition, and we include it as a comorbid condition for risk adjustment.

ACTING CHAIR HALL: Okay. So, what I am hearing is -- Jim, I'm sorry, do you still have a question?

Okay. So, what I am hearing is that there is a minor difference there, and we would like to know if the two developers can agree one way or another or if they feel that difference should be preserved. And hopefully, that response would get to Jeff's and some of the other concerns.

Okay. Any other larger -- Richard?

MEMBER BANKOWITZ: The question for NQF, now when we emerge, if we emerge with this harmonized model -- and I hope we do, and I think this is a great process, so I applaud it -- then who will be responsible for maintaining the model? Will it be a consortium of these two? Because we wouldn't want to see it sort of diverge as we go forward.

DR. BURSTIN: That is an excellent point. I think we would endorse both of those measures, and the expectation would be, as those measures come up for maintenance, we would again look to see, are they staying harmonized or is there any divergence?

But, again, this is pretty new territory for us. So, again, we are open to other suggestions if you think there is a better process.

MEMBER KELLY HALL: Yes, that was related to my intellectual property question.

ACTING CHAIR HALL: But, at the

same time, they are both sponsored by CMS. So, that is potentially in their favor.

MEMBER KELLY HALL: Right, right.

ACTING CHAIR HALL: Yes, Laurent?

MEMBER GLANCE: Just so I understand, in theory, it would be possible for both measure developers to respond to each one of these points and state that they feel strongly that there are reasons why they need to keep the measure as presented. Is that correct?

DR. BURSTIN: We are going to ask them to put forward a joint response where they are going to clarify each of those issues. They could justify and provide justification for why they think it needs to be different. But, again, they have already heard the stated preference of the Committee.

I'm sorry, one other point. I also think there are also opportunities to think about sort of short-term as well as longer-term harmonization here. So, I think

in the short-term here we are really saying there is an opportunity for harmonization, but we certainly heard from Yale that there was sort of potentially some interest in thinking about getting to the health plan level, potentially some interest in NCQA, again both sponsored by CMS. One question might be, the longer-term strategy I think, ideally, is this does become one measure usable as cascaded up and down.

I don't think we are there yet. That would be pretty radical changes to the measure, and neither developer is ready to do that. But I also think that might be a stated preference of the Committee in the longer-term as well.

ACTING CHAIR HALL: Thank you, Helen.

Jeff?

MEMBER GREENWALD: Just to reiterate a point, I am sorry, from early this morning, and for all the reasons that Helen

just brought up and some of the other concerns that have been brought up during the day, I wonder, given that this is an expedited review, whether we ought to think about an expedited re-review and not go the full distance at three years.

I understand that there are opportunities shorter than that to bring up concerns that would bring it back to review, but I think a structure process around that would be helpful, given the timescale that we have been talking about as well as the national implications of these numbers that we are dealing with.

MEMBER JHA: So, I am going to second that. And specifically given that neither of these, I mean both of these measures were initially voted down yesterday. Neither of them got super-majorities. I mean, they both won comfortably. So, there is enough concern on the Committee that I think a one-year checkup seems like a very

reasonable thing to do.

ACTING CHAIR HALL: I want to third that second. Great.

(Laughter.)

Thank you, Frank.

It has been clear through our discussions, obviously, that there is still a fair amount of discomfort on a number of aspects of this entire approach.

So, I don't see any other cards up. Going once, going twice.

So, I will turn to the NQF and ask how we move forward.

MR. AMIN: Okay. Thank you very much for leading that on the fly. It was a great job.

So, this was a great session. I think we got a lot of the preference of the Committee through this harmonization process.

So, we are going to ask for this response by the 13th. That is the timeline that we are dealing with. Again, we apologize

for the short turnaround time, particularly to the developers. But, as we discussed, it is an expedited review.

And I will turn it over to our leader here, Alexis, to give us some description on the next steps, and then we will go from there.

MS. FORMAN MORGAN: Sure. As Taroon stated, for the measure developers, we need your responses to the Committee's suggestions by next Tuesday, December 13th.

And then, we will send out an availability survey, so that the Committee can meet either next Thursday or next Friday via conference call. Then, we will go from there. We will have the final vote on endorsement, if needed, at that point, over SurveyMonkey.

MR. SAUNDERS: Could I ask a clarification question?

MS. FORMAN MORGAN: Yes.

MR. SAUNDERS: So, your expectation for us is a joint response to

these questions, and that is our deliverable for Tuesday? We are not delivering to you a full implementation of the specification? We are not presenting to you the rerun weights and everything? Okay.

ACTING CHAIR HALL: That is correct. We are looking for a response that says, "We believe we could adopt the same approach on this item" or "We would like to justify and explain why we have a different approach on this item."

PARTICIPANT: Unless you want to run all those models by next Tuesday, and we would be happy to look at those.

(Laughter.)

ACTING CHAIR HALL: CMS in the room, is there a bank account open right now?

(Laughter.)

While they are talking, Jim?

MEMBER BELLOWS: Oh, I don't know procedurally what is the right time, but I just wanted to thank NQF and the Committee.

If this harmonization goes as planned, I think it will be just a tremendous win for this area that I personally did not anticipate walking into this process. So, thank you.

MS. FORMAN MORGAN: Thank you.

So, we are now at the end of our two-day meeting.

And staff would like to thank all of you for your great participation.

We do have lunch available for you all. If you would like to stay here and eat it, that is perfectly fine. It is in the back room. Or if you would like to take it with you, it is boxed.

So, once again, we thank you for this two-day meeting, and we will see you next week via conference call.

ACTING CHAIR HALL: And on behalf of Eliot and Sherrie, I want to thank everybody as well.

MEMBER LAZAR: Yes, this is Eliot.

First of all, I want to obviously

thank all of you, thank Sherrie, and particularly thank Bruce Hall for stepping in without any preparation whatsoever and handling the finish with real aplomb.

(Whereupon, at 12:24 p.m., the meeting was adjourned.)

C

C E R T I F I C A T E

This is to certify that the foregoing transcript

In the matter of: Readmissions Steering Committee

Before: NQF

Date: 12-06-11

Place: Washington, DC

was duly recorded and accurately transcribed under my direction; further, that said transcript is a true and accurate record of the proceedings.

Neal R. Gross

Court Reporter